Investigation into the twin-arginine translocation pathway of halophilic and thermophilic archaea

Kwan, Daniel

Award date:
2009

Awarding institution:
University of Bath

Investigation into the Twin-Arginine translocation pathway of halophilic and thermophilic archaea

Daniel Kwan

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Pharmacy & Pharmacology

July 2009

TABLE OF CONTENTS

LIST OF FIGURES

List of Tables

Acknowledgements

Firstly, I would like to thank Albert Bolhuis for being such an impressive supervisor. Not only for his wide range of scientific knowledge, useful input and dedication to this project but also for providing support during the harder times.

I would particularly like to thank Michael Danson (Chair) for his useful discussions on my project during the Extremophile talks. Also thanks to everyone from Bath University, Department of Pharmacy & Pharmacology. They made me feel welcome during the transition to this University, especially the "Flying Boots" football team. They also continue to make this Department an enjoyable environment to work in.

I would also like to thank Colin Robinson (Previous chair and advisory committee) for his helpful advice during the first two years of this project. Many thanks as well to the rest of the Warwick Biology Department for all their encouragement and help. I would also like to acknowledge the BBSRC for their financial backing. I also appreciate the help provided by Sonja Albers and her group. They kindly allowed me to work with them at the University of Groningen and taught me several *Sulfolobus* techniques such as creating knockouts and performing protein purification.

Special thanks go to Judith Thomas who during the first two years supported me and helped me get to grips with working in a Microbiology lab, she still continues to be a great friend. Finally I want to express my gratitude to my Mum and Dad and also Natalie Meade, all of whom have continually believed in me and put up with me even when things were hard.

Declaration

The work presented in this thesis is original work conducted by myself under the supervision of Dr. Albert Bolhuis. The research was funded by a BBSRC studentship. All sources of information have been acknowledged by means of references. None of this work has been used in any previous application for a degree at this or any other University or institute of learning.

Abstract

The Twin arginine translocation pathway translocates fully folded proteins across cellular membranes and is only utilised by proteins that fold before translocation. It is a unique process that is found in many bacteria, archaea and also in plant chloroplasts. Investigation of the bacterial and thylakoidal systems has revealed much of the substrates and the components involved in their translocation. Unfortunately, there are still many unanswered questions such as how substrates are directed to the membrane and the actual mechanism of translocation.

This thesis specifically investigates the Tat pathway of halophilic and thermophilic archaea. To date, there has been a lack of research into the archaeal Tat pathway and it is possible that there are unique adaptations because of the extreme environments that these organisms inhabit. Chapter 3 specifically investigates the thermophiles *Sulfolobus solfataricus* and *Sulfolobus tokodaii* and attempts to purify their Tat complexes. By doing so it was hoped to learn more about the Tat components and their interactions. Further experiments were also performed to determine if the two *S. solfataricus* Tat operons provide specificity to the Tat substrates that translocate.

Four separate areas of the Tat pathway of halophilic archaea (haloarchaea) were investigated in Chapters 4-7. Firstly, site-directed mutagenesis was used to analyse the signal peptides of haloarchaeal Tat substrates in more detail. Consequently, the resulting data led to the use of bioinformatics to analyse the Haloarchaeal signal peptide. The bioenergetics of the Tat system was then determined by analysing the effect of a variety of ionophores on translocation of the Tat substrates AmyH and SptA. Finally, a series of folding and stability assays were used to increase our understanding of AmyH, which could provide further information on why this protein, like many other haloarchaeal proteins, requires the Tat pathway for translocation.

Abbreviations

AmyH	Amylase from *H.hispanica*
APS	Ammonium persulphate
Ara	Arabinose
ATP	Adenosine triphosphate
B. subtilis	*Bacillus subtili*s
Bis-acrylamide	N',N'-methylene-bisacrylamide
Brock medium	Brocks salt basal medium
BSA	Bovine serum albumin
CCCP	carbonyl cyanide m-chlorophenylhydrazone
Cfu	Colony forming unit
Ci	Curie
C-terminal	carboxy-terminal
DDM	n-Dodecyl-β-D-maltoside ULTROL Grade
DHFR	dihydrofolate reductase
DMSO	Dimethyl sulfoxide
DMSOr	DMSO reductase
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	ethylenediaminetetraacetic acid (sodium salt)
F	Phenylalanine
GFP	Green florescent protein
Haloarchaea	halophilic archaea
His	Histidine
H. hispanica	*Haloarcula hispanica*
H. marismortui	*Haloarcula marismortui*
H. mediterranei	*Haloferax mediterranei*
H. salinarum	*Halobacterium salinarum*
H. volcanii	*Haloferax volcanii*

Hydrophob.	Hydrophobicity
IMV	Inverted membrane vesicles
IPTG	Isopropyl β-D-thiogalactopyranoside
K	Lysine
kBp	Kilo base pair (s)
KCl	Potassium chloride
kDa	Kilodalton(s)
L	Leucine
LB	Luria-Bertani Medium
Lep	Leader peptidase
M	Molar
Mins	Minutes
Mo1	Molybdopterin oxidoreductase
MIC	Minimum inhibitory concentration
mRNA	Messenger RNA
NaCl	Sodium Chloride
N. amylolyticus	*Natronococcus amylolyticus*
N. pharaonis	*Natronomonas pharaonis*
N. occultus	*Natronococcus occultus*
N-terminal	Amino terminal
OD	Optical density
Od1	hypothetical protein (oxidoreductase)
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
phoD	Phosphodiesterase
PPO	2,5-diphenyloxazole
Pre	Precursor
PMF	proton motive force
Qo2	Quinol oxidase-2
R	Arginine
RR	Two Arginines
S	Serine
S. acidocaldarius	*Sulfolobus acidocaldarius*
SDS	Sodium dodecyl sulphate

Sec pathway	general secretory pathway
Sig.	Signal
SMF	Sodium motive force
SoxL-1	Rieske iron-sulfur protein-1
SoxL-2	Rieske iron-sulfur protein-2
spp.	Species
SRP	Signal recognition particle
S. solfataricus	*Sulfolobus solfataricus*
S. tokodaii	*Sulfolobus tokodaii*
T	Threonine
Tat	Twin arginine translocase
TorA	trimethylamine N-oxide reductase
Tris	N-tris(hydroxymethyl)methylamine glycine; N-(2-hydroxy-1,1-bis(hydroxymethyl)ethyl) glycine
tRNA	transfer RNA
UV	Ultra violet
V	Valine
w/v	weight per volume
x	Any amino acid
ΔΨ	Electrical gradient
ΔpH	Chemical gradient
Δ(x)	Mutant strain lacking the (x) gene

Chapter 1: Protein transport through the Twin Arginine translocation pathway

1.1.1 Introduction

Protein translocation across cellular membranes is an essential process in all three domains of life. Knowledge into this area is not only of scientific interest but also has huge potential in the biotechnological industry [1]. Due to the hydrophilic nature of proteins, they are unable to passively cross the hydrophobic membrane without the aid of proteinaceous channels. As a consequence, a variety of protein transport pathways have evolved. In prokaryotes there are two main pathways for the translocation of proteins across the cytoplasmic membrane. The first is the Sec pathway which translocates the majority of these proteins. For this reason, the Sec system has already been extensively researched. The second pathway is the Twin Arginine Translocation (Tat) pathway, which is the focus of this thesis. The unique feature of the Tat pathway is that it specifically translocates proteins that are in a fully folded state, although little is known about this process.

1.2 The Sec pathway

1.2.1 Overview of the Sec pathway

The Sec pathway is conserved in all three domains of life and allows unfolded proteins to translocate across the membrane [2]. There are two modes of Sec-dependent translocation, being either co-translational or post translational. Eukaryotes mainly use the co-translational method where the substrate translocates through the membrane while it is translated using a process known as the signal recognition particle (SRP) pathway [3, 4]. The substrate emerges from the ribosome allowing the SRP to bind which will arrests protein synthesis (Figure 1.1) [2]. In mammalian cells, SRP consists of a 7S RNA molecule and 6 proteins: SRP9, SRP14, SRP19, SRP54, SRP68 and SRP72 [5]. After binding, the entire complex is targeted to the ER membrane with the aid of the SRP receptors SRα and SRβ. The ribosome then docks with the translocation channel and the SRP is released. This allows translation to continue with the ribosome pushing the protein through the Sec channel as it is synthesised. The channel is made up of three proteins, Sec61α, Sec61β and Sec61γ as shown in Figure 1.1 [6].

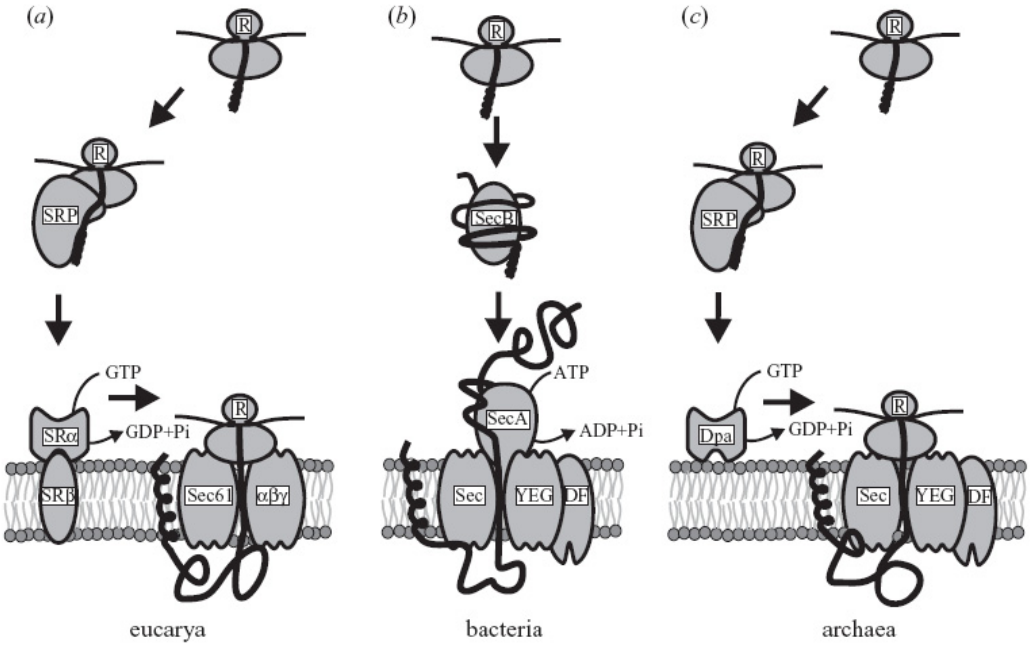

Figure 1.1 The Sec translocation pathway of eucarya, bacteria and archaea. The Sec pathway of a) eucarya, b) bacteria and c) archaea. Only the main routes of the pathway are shown and not the bacterial SRP route or a possible post-translational route of archaea. R represents the ribosome. This figure was taken from Bolhuis., 2004 [2].

Bacteria mainly use the post-translational method, which involves the substrate first being synthesised fully in the cytoplasm. A chaperone such as SecB then targets the substrate to the membrane and also prevents the substrate from folding (Figure 1.1) [7, 8]. At the membrane the substrate interacts with the ATPase SecA which provides the driving force for the pathway [9]. Translocation occurs through a protein channel formed by SecY, SecE and SecG, which are homologous of Sec61α, γ and β respectively (Figure 1.1). For successful translocation, the substrate requires several cycles of SecA binding, membrane insertion and de-insertion [10]. Also present are the accessory proteins SecD and SecF which have been proposed to play a role in the cycling of SecA. Bacteria also contain an SRP pathway which is specifically used for the insertion of membrane proteins such as YidC [11]. The bacterial SRP complex is a simpler version compared to the eucaryal system and just consists of 4.5S RNA and Ffh (**F**ifty **f**our **h**omologue) protein; the latter is a homologue of SRP54. After the ribosome binds the chaperone, FtsY assists the complex to the membrane for translocation [12].

It has been speculated that the reason eukaryotes translocate co-translationally and prokaryotes post-translationally could be related to protein folding which is intimately linked to protein translocation [13]. Eukaryotic proteins are larger than prokaryotic proteins and also have the tendency to contain multiple domains [14]. They are also translated at a slower rate then bacterial proteins (2-5 fold slower) [15]. These factors produce the requirement for co-translational folding which in-turn could also mean that co-translational translocation is needed. Prokaryotic proteins are shorter and are translated faster which probably allows for the post-translational translocation.

Archaea form the third domain of life and are explained in more detail in Section 1.4. They do possess an SRP pathway and evidence suggests that it includes aspects from both the bacterial and the eukaryal pathways [16]. The SRP consists of a 7S RNA and protein homologues of SRP54 and SRP19 [16]. A homologue to the eukaryotic SRα has also been found (Dpa) which is likely to direct the ribosome-nascent chain to the membrane (Figure 1.1). However, there appears to be no homologue of SRβ. After binding to the translocation channel, the SRP will be released allowing the substrate to translocate as it is synthesised. It appears that the Sec translocation complex consists of homologues to SecY/Sec61α, SecE/Secγ and SecG/Secβ [17]. Finally, no SecA homologue has been found in archaeal species, although it is important not to overlook the possibility that an adapted version of the bacterial post-translational system could still exist.

1.3 The Tat pathway in bacteria/chloroplasts

1.3.1 Introduction to the Tat pathway

The Tat pathway is not ubiquitous to all organisms but can be found in many bacteria, archaea and also in plant chloroplasts (in which it is also known as the ΔpH-dependent pathway). It specifically allow the translocation of fully folded proteins across the membrane [18, 19]. Although folding seems to be a prerequisite, in the thylakoidal system there is evidence that some Tat substrates do not have to fold correctly to use the pathway [20]. In bacteria less then 10% of secretory proteins use the Tat pathway and appear to do so only if they fold in the cytoplasm [21]. Evidence suggests that this is usually because the Tat substrate has to bind complex co-factors that have to be inserted enzymatically in the cytoplasm, although it is also possible that some proteins simply fold very rapidly [22, 23].

Tat dependent proteins tend to be involved in important processes such as the respiratory and photosynthetic energy metabolism [24]. In the Gram-negative bacterium *Escherichia coli* the Tat pathway is only essential for anaerobic respiration as it translocates a number of important redox-binding proteins [25]. Even though the Tat pathway is not essential, it is expressed constitutively under all growth conditions [26]. If blocked then cells grow in short chains and can no longer survive anaerobically [27]. The largest *E. coli* Tat substrate is the 142 kDa FdnGH subcomplex of formate dehydrogenase-N [28] and the smallest is the 20 kDa formate-dependent nitrite reductase (NrFc) [29, 30]. This large difference in size must mean that the pathway has a method to dramatically change the size of the pore without losing membrane integrity during translocation [20].

1.3.2 Signal peptides of Tat and Sec substrates

Targeting of Tat and Sec substrates to their respective pathways requires a short stretch of amino acids found at their N-terminal. These amino acids form a so-called signal peptide consisting of three domains: a charged N-domain, a hydrophobic H-domain and a polar C-domain (Figure 1.2) [31]. There are slight differences between

Sec and Tat signal peptides that allow the substrate to be directed to the correct pathway. Firstly, the N-domain of the Tat signal peptide tends to be longer and the H-domain less hydrophobic then that of the Sec signal peptide [32, 33]. Also in chloroplasts there are positively charged residues at the end of the C-domain that appear to contribute to a "Sec-avoidance" system [34].

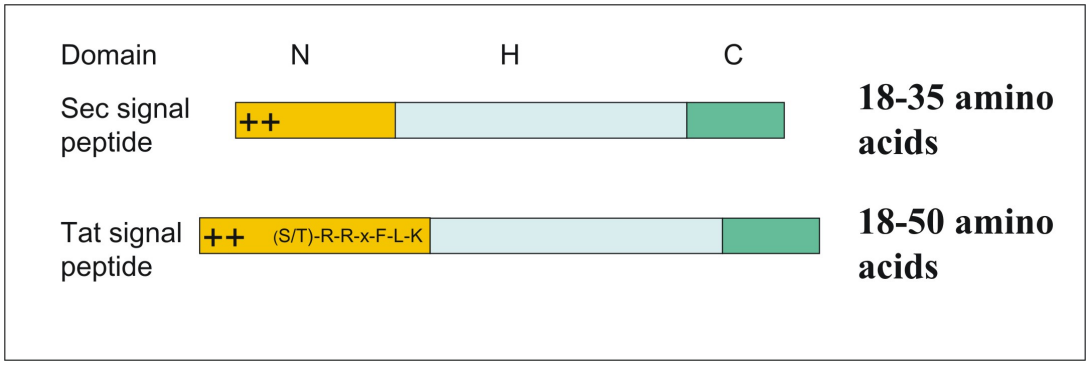

Figure 1.2 The Tat and Sec signal peptide. The 3 domains of the Tat and Sec signal peptide. The N-domain (Orange) is the positive domain. The H-domain (Blue) is the hydrophobic domain and the C-domain (Green) is the polar domain.

The main distinguishing feature of the Tat signal peptide is the presence of a conserved motif, (S/T)-R-R-x-F-L-K within the N-domain (Figure 1.2) [35]. The two arginines (RR) have been found in nearly all known Tat substrates and the other five amino acids occur at a frequency greater than 50% [22]. The arginines play an important role for the targeting of proteins to the Tat pathway. For example in *E. coli,* if the second arginine (RR->RK) is mutated to a lysine then translocation is completely blocked [32, 36]. If the first arginine is changed to a lysine (RR->KR) some translocation does occur but efficiency is greatly reduced [37]. Arginine and lysine are both positively charged molecules, however lysine is shorter and is also not branched. Since mutation of the arginines to lysines (RR->KR->RK->KK) blocks or greatly reduces translocation, this suggests that these mutations reduce the specific interactions with the Tat recognition site which is most likely located on TatB or TatC [38, 39].

There have been several cases where a Tat substrate is a heteromeric complex of two or three subunits where only one of these subunits contains a signal peptide [40, 41]. For example, the hydrogenase-2 from *E. coli* is composed of a small subunit HybO, and a

larger subunit HybC. Only HybO contains a signal peptide but it forms a complex with HybC in the cytoplasm allowing both proteins to translocate via the Tat pathway [42]. This system has been called the Hitchhiker co-translocation mechanism and provides further evidence for the requirement of Tat substrates to fold correctly prior to translocation [43].

The signal peptide appears to be the main requirement for proteins to utilise the Tat pathway. For example, Green fluorescent protein (GFP) found in jellyfish (normally not a protein that is transported across a membrane) will translocate across the membrane of *E. coli* if the trimethylamine N-oxide reductase (TorA) Tat signal peptide is attached to its N-terminus [44]. Once a Tat substrate has translocated the signal peptide is removed by a type I Signal Peptidase, which in *E. coli* is called LepB [32, 45].

1.3.3 Components of the Tat complex

E. coli possesses four genes encoding Tat translocation components. *TatA, tatB* and *tatC* form an operon while *tatE* is located elsewhere (Figure 1.3). There is also a *tatD* gene known to encode a cytoplasmic protein that is not involved in the pathway but analysis from Wexler *et al*., (2001) [46] indicated that it was expressed within the same operon. RT-PCR was performed by Weiner *et al*., (1998) [29] who confirm that *tatD* was part of the *tatA* operon however it has been suggested that their results were actually false positives. Wexler *et al*., (2001) [46] continued their work by performing a Northern blot using a variety of probes on the predicted operon. Their conclusion was that *tatD* was part of the same operon however it appeared to be transcribed at lower levels than *tatABC* under the conditions tested. Interestingly, it is not uncommon in *E. coli* for genes encoding components of the protein secretion apparatus to form operons with genes of unrelated function. Contradictory to these findings, recent results published by Matos *et al*., (2009) [47] suggest that TatD is actually involved in the quality control of the Tat pathway. They came to this conclusion since the Tat substrates NrfC and NapG lacking their FeS centres are rapidly degraded in wild-type cells but remain stable in a Δ*tatD* strain.

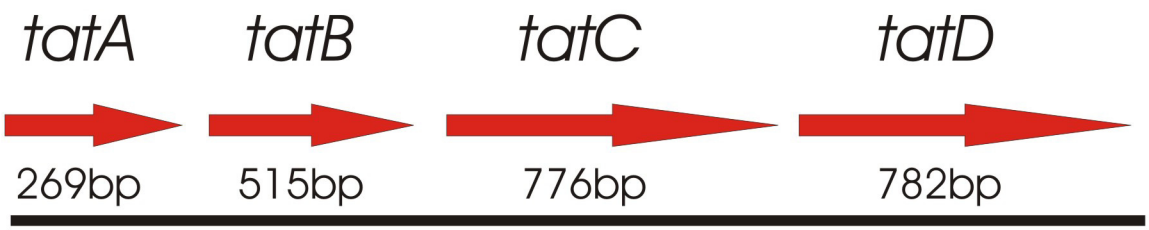

Figure 1.3 The *E. coli's* Tat operon. The genes that make up the *E. coli* tat operon and their individual sizes. TatE (Not shown) is 203bp in size.

In *E. coli*, TatA (~9.6 kDa) and TatE (~ 7kDa) are homologous with almost 60% identity [48, 49]. They consist of a hydrophobic transmembrane α-helix followed by a amphipathic α-helix on the cytoplasmic side of the membrane (Figure 1.4) [50]. *TatE* is thought to be a cryptic gene duplicate since it is expressed approximately hundred-fold less then TatA [26]. In *E. coli*, TatA is not essential since TatE can complement its loss but translocation is greatly reduced. If both proteins are absent then the Tat pathway can no longer function [51].

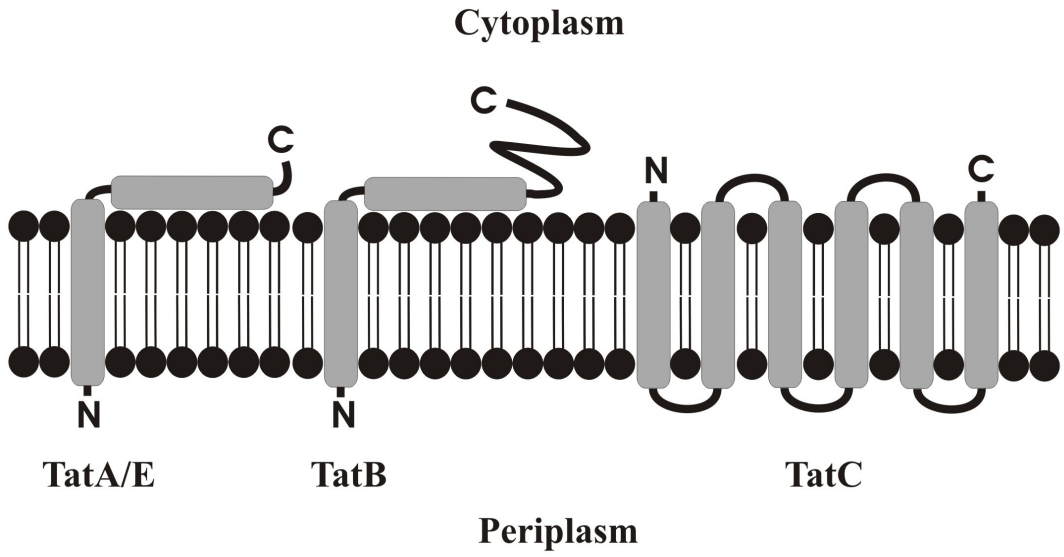

Figure 1.4 The Tat components of the *E. coli* Tat pathway. The 4 different *E. coli* tat components and their predicted interactions with the cell membrane.

TatB (~18.4 kDa) also contains a hydrophobic and an amphipathic domain but it has a longer C-terminal compared to TatA (Figure 1.4) [52]. Although the proteins are similar in structure there has only ever been found to be approximately 25% amino acid identity between these proteins [48]. The only conserved amino acid when

comparing all bacterial TatA and TatB proteins is a glycine found at the carboxy-terminal end of the predicted helix. TatA also contains a conserved phenylalanine immediately before this glycine while a conserved proline has been found to be directly after in TatB proteins [23]. For Gram negative bacteria both TatA (or TatE) and TatB are essential for translocation, which shows that TatB has a function that is distinct from TatA or TatE [53].

TatC (~28.9 kDa) is a much larger protein that consists of six hydrophobic domains. Each domain spans the membrane with the N- and C-terminus predicted to be on the cytoplasmic side (Figure 1.4) [51]. It appears that TatC is essential for the Tat pathway to function [51]. Purification of the *E. coli* TatABC complex revealed that in detergent it was approximately 600 kDa with TatB and TatC present at a 1:1 ratio [39, 54]. TatA is expressed approximately 20-fold higher and mainly forms separate homo-oligomeric complexes [26, 55]. In *E. coli,* blue native gels revealed that these TatA complexes range from 65 kDa to 650 kDa with the majority being around 400 kDa [56].

The chloroplasts ΔpH-dependent pathway translocates proteins through a complex across the thylakoid membrane. This complex consists of the proteins Tha4, Hcf106 and cpTatC which are homologues of TatA, TatB and TatC respectively [57].

1.3.4 The driving force of the Tat pathway

Unlike the Sec pathway, ATP hydrolysis does not provide the driving force for Tat translocation. Instead, the process in bacteria is driven completely by the proton motive force (PMF) which is formed by the combination of both the electrical ($\Delta\Psi$) and the chemical gradient (ΔpH) (Figure 1.5) [58]. The electrical gradient is produced by the difference in charge on either side of the membrane while the chemical gradient is determined by the difference in pH [59]. It is not yet known if both the electrical and the chemical gradient are involved in driving the Tat system [45].

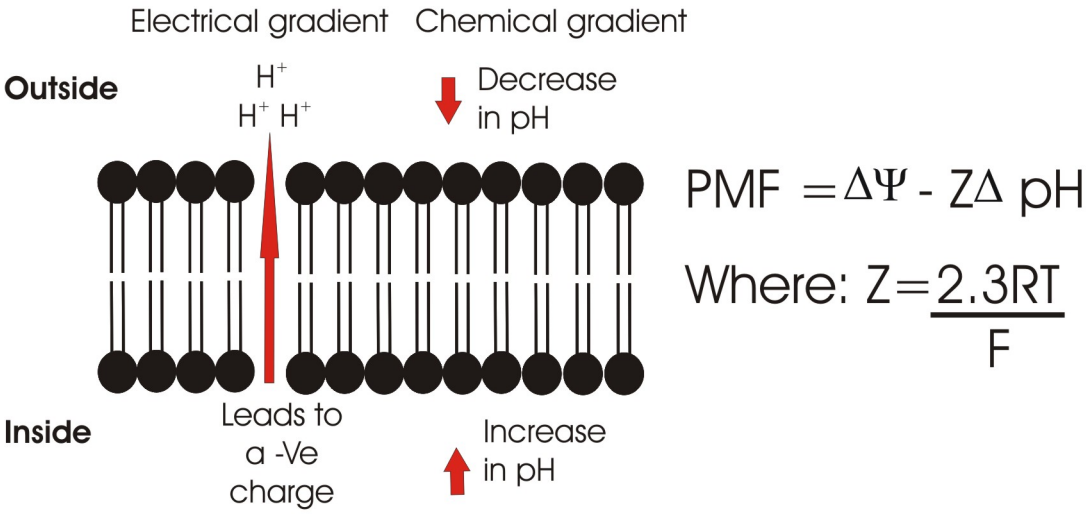

Figure 1.5 The proton motive force. Protons pumped across the membrane through the respiratory chain leads to a change in charge distribution and concentration of protons. This generates an electrical gradient with a negative charge on the cytoplasmic face (inside) and a positive charge on the trans side of the membrane (outside). The change in proton concentration leads to a difference in pH across the membrane known as the chemical gradient. The electrical and chemical gradients are not always necessarily linked but are both contributing factors to the PMF as shown in the above equation. R = the gas constant, T = the absolute temperature and F = Faraday constant.

1.3.5 Translocation of Tat substrates in Gram negative bacteria

In Gram negative bacteria, the Tat pathway allows proteins to translocate across the inner membrane into the periplasm. If the protein is then required to translocate across the outer membrane then this is done by a separate pathway [60]. For example, the *E. coli* H7 flagellin is a Tat substrate that then translocates through the outer membrane by a type III protein transport mechanism [61]. How the Tat pathway functions is still not clear but the current hypothesis is described in Figure 1.6.

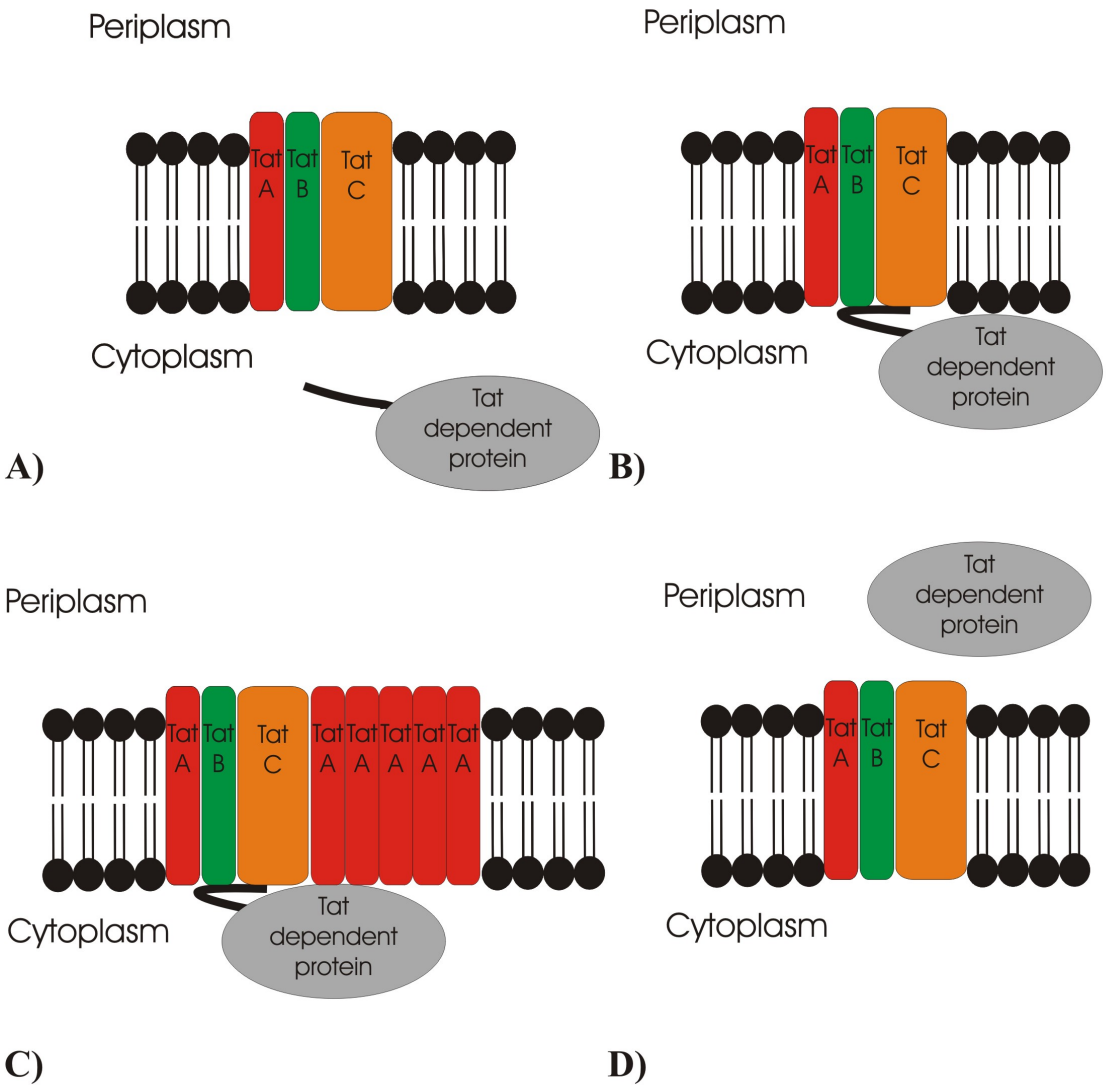

Figure 1.6. Proposed mechanism for Tat translocation in Gram negative bacteria. Stage A shows the situation when the Tat substrate is first translated and folded. Stage B shows the Tat substrate interacting with TatB and TatC. Stage C shows the incorporation of the varying sized TatA complexes into the membrane. Stage D shows the Tat substrate after translocation has occurred with the signal peptide cleaved and the removal of TatA complex.

It is thought that the Tat substrate first translates and folds in the cytoplasm (Figure1.6A), and then interacts with the TatABC complex in an energy independent step [38]. It is not known if the targeting is aided by chaperones although research has shown that translocation efficiency of the Tat substrate TorA is greatly enhanced in the presence of the cytoplasmic chaperone TorD. The presence of TorD does not however improve the translocation of other Tat substrates and thus is specific to TorA. Cross linking has shown that TatB interacts with the hydrophobic core of the signal peptide and the substrate itself, while TatC interacts solely with the twin

arginine motif (Figure 1.6B) [38, 39]. Once these interactions have occurred then translocation will begin.

It is conceivable that the differently sized TatA complexes allow the pore to vary depending on the protein that is translocating. This will allow the pore to fit tightly around the substrate so that membrane integrity can be maintained [48]. Since TatB and TatC form large complexes it is unlikely that they act solely as receptors. Recent results from Jakob *et al*., (2008) [62] appear to corroborate this hypothesis. Thus it seems likely that TatB and TatC also form the core of the pore, which then varies depending on the amount of TatA that binds (Figure 1.6C) [48]. It has been found that in thylakoids, these TatA complexes only associate when both the substrate and the proton gradient are present. If one of these factors is lost then the majority of TatA will disassociate again [63].

As mentioned previously, the method of translocation is still unclear and at present there are three main hypotheses. Natale *et al*., (2008) [64] suggest that after substrate binding and the recruitment of TatA, the TatB/C complex undergoes a conformational change which pulls the Tat substrate through the membrane. It is thought that TatA weakens the membrane to allow this process to occur. This of course will lead to a short lived hole in the membrane which may account for the large proton flux seen with Tat transport [65]. Gouffi *et al*., (2004) [63] propose that the *E. coli* TatA has a dual topology since they were able to detect the C-terminal amphipathic helix on both the cytoplasmic and the periplasmic side of the membrane. It has therefore been suggested that the amphipathic helix of TatA "flips", transferring the substrate across the membrane without affecting membrane integrity. Finally, Gohlke *et al*., (2005) [66] used electron microscopy and random conical tilt reconstruction on TatA which suggests that the protein is in fact closed by a lid at one end which may gate access to the channel allowing transport by diffusion. It must be remembered that these hypothesis are not mutually exclusive, it is possible that a combination of these features are utilised for translocation.

During or shortly after translocation has occurred the signal peptide is cleaved by the signal peptidase LepB, and the TatA complex dissociates from TatBC (Figure 1.6D) [45].

1.3.6 Translocation of Tat substrates within Gram Positive bacteria

The Tat pathway in Gram positive bacteria is likely to function similarly to that of the Gram negative system (Section 1.3.5). The main difference is that no TatB protein has yet been found, with the exception of the Tat system in *Streptomyces* spp. [67]. It has been suggested that TatA has adapted to also perform TatB's function due to the similarities between the two proteins [68]. This is backed up by the findings that an *E. coli* Δ*tatB* mutant still has a functioning Tat pathway when there are specific mutations in TatA. These mutations are confined to the five amino acids found between the periplasmic and transmembrane domains [68]. Interestingly, the mutations tended to replace the amino acid with an aspartate which is also present at a similar position in TatB, see Figure 1.7. This proposed dual function of Gram positive TatA proteins has been proven for the TatAd from *B. subtilis* which is capable of complementing the Tat pathway of both a Δ*tatAE* and a Δ*tatB E. coli* mutant [69].

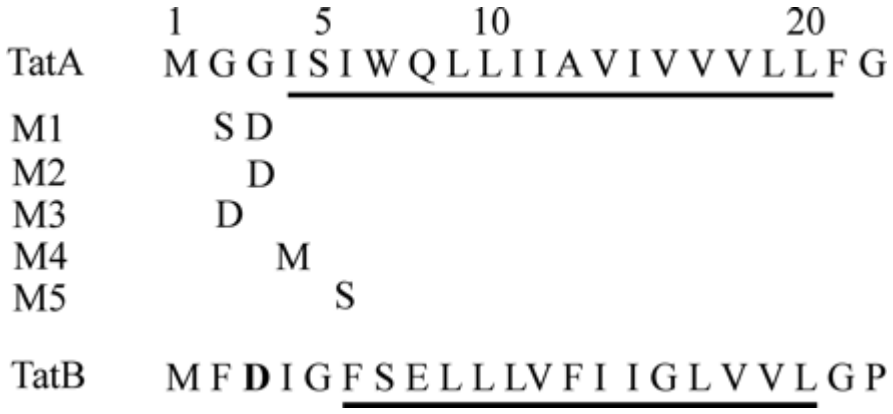

Figure 1.7. The 5 mutations in TatA that allowed a functioning Tat pathway in an *E. coli* Δ*tatB* mutant. This diagram is from Blaudeck et al., 2005 [68]. TatA is the primary sequence corresponding to the short amino-terminal periplasmic region and the proposed transmembrane domain (Underlined) of TatA. M1 to M5 are the mutant TatA proteins with alterations also stated. The corresponding homologous region of TatB is also shown for comparison.

The other major difference is that the Gram positive *Bacillus subtilis* has been found to express multiple TatA and TatC proteins [70]. Research revealed that it contains two Tat operons that encode a TatA and a TatC gene. The first complex (denoted $TatA_dC_d$) is only transcribed under phosphate starvation to specifically allow the translocation of a phosphodiesterase (PhoD) [71]. Removal of this operon will stop PhoD from

translocating [70, 72]. The second complex (denoted $TatA_yC_y$) cannot translocate PhoD but instead translocates another *B. subtilis* Tat dependent protein, YwbN [73]. Recently, two other *B. subtilis* substrates, QcrA and YkuE have been found although it is not yet known if they have a preferred Tat complex [74, 75].

In several Gram positive bacteria, TatA has been found to form micelles in the cytoplasm [76, 77]. It has been suggested that this is because TatA also acts as a chaperone. The micelle will interact with the signal peptide and transport the substrate to the Tat complex [78]. Once reaching the complex, TatA could then interact with the membrane to form the pore. However, this theory is still heavily debated since the over expression of TatA in these experiments may have led to formation of these micelles in the cytoplasm [79].

1.4 The Archaeal Domain

1.4.1 Introduction into the Archaeal domain

Over the past 100 years, phylogenetics have been based on Chatton's proposal of a eukaryote-prokaryote system where life is divided into these two separate categories [80]. As molecular and cytological understanding of the cell improved it was realised that this current distinction was not accurate enough. This is because a group of organisms, the archaebacteria represented prokaryotes on a cytological level but were different to both domains when analysed on a molecular level [81]. It was therefore realised that the current two-domain phylogenetic system needed to be restructured.

Woese *et al*., (1990) [82] devised a new system that looked at the small subunit of the ribosomal RNA. From this they realised that organisms could be grouped into three distinct domains known as the Bacterial, Eukaryotic and Archaeal domains. By looking more closely at the sequencing it was determined that all three groups had evolved from a single ancestor and it was possible to accurately create the phylogentic tree shown in Figure 1.8.

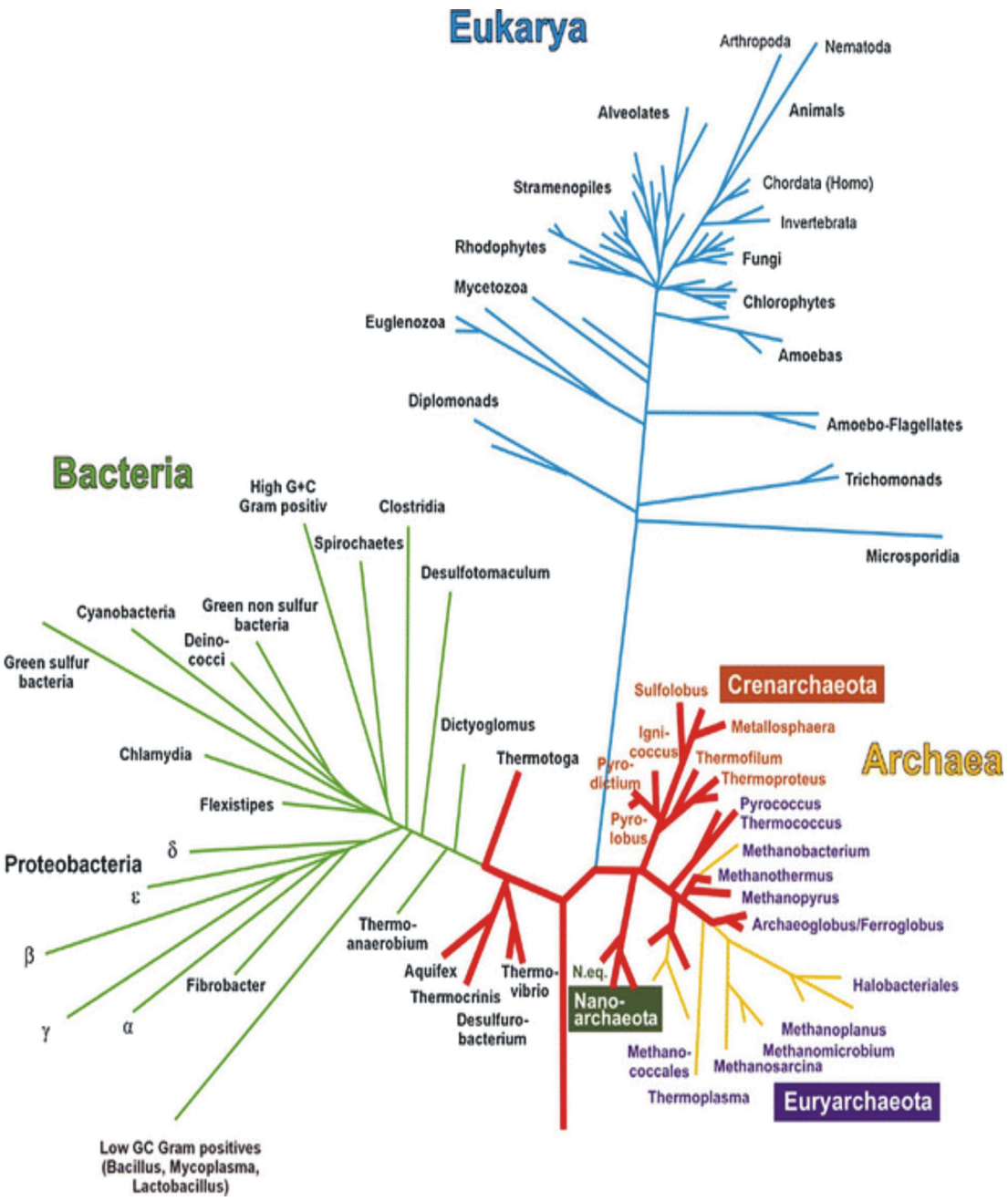

Figure 1.8 The Phlogenetic Tree of Life. The Phylogenetic Tree of Life is divided into three domains that stem from a common ancestor. The Bacterial domain is shown by the green lines and red lines that branch left from the origin, the Eukaryotic domain is shown by the blue lines. The Archaeal domain is split into 3 groups, Crenarchaeota and Nanoarchaeota are shown by the red lines that branch to the right from the origin (the thermophiles) while Euryarchaeota are shown by the yellow lines. This image was taken from Stetter *et al*., (2006) [83] who suggests that the origin of life is a thermophile.

As shown by the phylogenetic tree, archaea are divided into two major kingdoms, the Crenarchaeota and the Euryarchaeota. The Euryarchaeota consists of the methanogens, exteme halophiles, thermophilic sulphate reducers, the *Thermoplasma* group and also the *Thermococcus-Pyrococcus* group [2]. The Crenarchaeota were

thought to only consist of sulphur-metabolizing organisms, however several other archaea have now also been included (This includes several mesophilic organisms) [84]. Finally, there is also the smaller nanoarchaeota kingdom which at present only contains the organism, *Nanoarchaeum equitans* [85].

However, the phylogenetic tree is still heavily debated as recent analysis on protein translation elongation factors suggest that this tree is inaccurate [86]. This evidence is supported to a lesser extent by the analysis of the SSU rRNA sequences [87] and the identification of an 11 amino acid insertion in the GTPase domain of the elongation factor 1 [88]. From these results it has been proposed that the ‘‘eocyte hypothesis,’’ is more realistic. Instead of the eukaryotes and archaebacteria being separate groups that share a common ancestor to the exclusion of eubacteria, the eukaryotes actually originate within the archaebacteria and share a common ancestor with *Crenarchaeota* (See Figure 1.9).

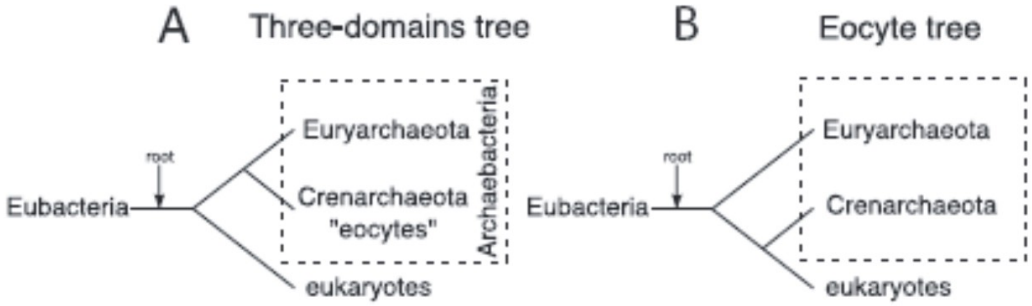

Figure 1.9 The eocyte hypothesis. *A*) The original 3-domain Phlogenetic tree, B) An alternative hypothesis, the eocyte tree where eukaryotes are suggested to originate within the archaebacteria. This diagram was taken from Cox *et al*., (2008) [89]

It has been found that the majority of archaea characterised grow in extreme environments such as high temperatures or pH. However, these extremophiles are most likely the minority, as metagenomic techniques have identified that archaea are very abundant in more moderate environments such as the oceans and soil [90].

Extremophiles have created a particular scientific interest because these organisms tend to have unique cellular adaptations due to the extreme environments they live in. For example the chemical composition of the cytoplasmic membrane is unique in

archaea. It is composed of lipids made of two phytanyl chains linked to a glycerol or other alcohol *via* an ether bond rather then an ester bond as found in bacteria and eucarya. The acyl chain is usually a fully saturated isoprenoid (Figure 1.10) [91]. One advantage of these adaptations is that it allows the membrane of thermophilic species to be more stable at high temperatures.

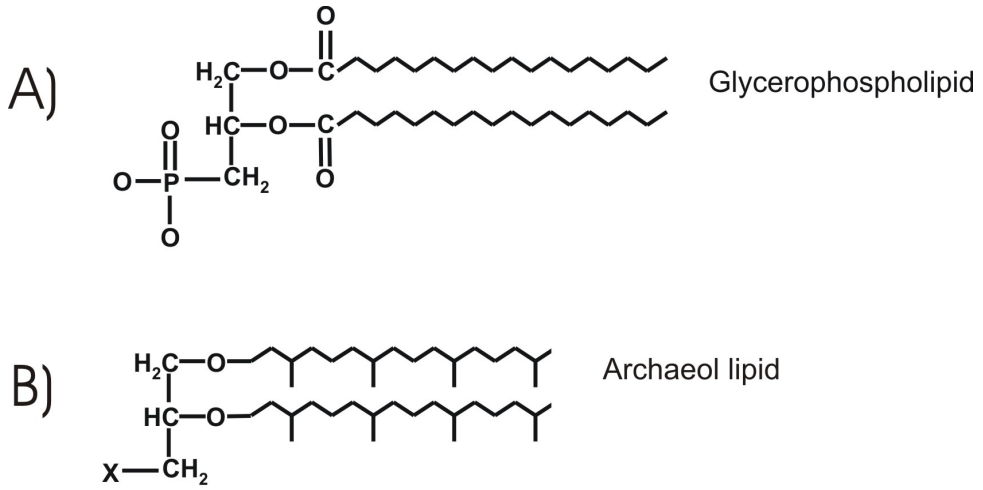

Figure 1.10 Chemical representation of the bacterial and archaeal membrane lipid. A) The basic lipid found in a bacterial membrane. B) The basic Archaeal membrane lipid which is two phytanyl chains linked to a glycerol by an ether bond. X represents the polar head group which in archaeol lipids is generally either a phosphoethanolamine, phosphoserine, phospho-myo-inositol, or a disaccharide [92].

Yen *et al*., (2002) [93] have found that a variety of archaeal species do utilise the Tat pathway. In fact 10 out of the 16 available archaeal genomes contain at least one *tatC* gene, as shown in Table 1.1 [94].

In extremophiles, all proteins and the cellular processes they are involved in need to be able to function in adverse conditions. It is therefore likely that the Tat pathway has had to adapt to be able to continue to function in many archaea [95]. This thesis investigates the Tat pathway of two very different groups of archaea, being *Sulfolobus* spp. and halophilic archaea (haloarchaea) which are both described in more detail in Section 1.5 and 1.6 respectively.

	Tat Components		
Archaea	**TatA/E**	**TatB**	**TatC**
Halobacterium sp.NRC-1	1	-	2
Pyrobaculum aerophilum	1	-	1
Archaeoglobus fulgidus	2	-	2
Aeropyrum pernix	2	-	1
Methanosarcina mazei Goe1	2	-	2
Pyrococcus horikoshii	-	-	-
Methanosarcina acetivorans str.C2A	2	-	2
Sulfolobus solfataricus	3	-	2
Sulfolobus tokodaii	2	-	1
Pyrococcus furiosus DSM3638	-	-	-
Pyrococcus abyssi	-	-	-
Thermoplasma acidophilum	1	-	1
Thermoplasma volcanium	1	-	1
Methanothermobacter thermautotrophicus	-	-	-
Methanocaldococcus janaschii	-	-	-
Methanopyrus kandleri AV19	1	-	-

Table 1.1 The Tat components that have been found in Archaeal species. This information is from Pohlschroder *et al*., 2004 [94]. Table shows the 16 sequenced Archaeal genomes and the Tat components that are present for each. – means that no homologue of that particular Tat component has been found in that genome.

1.5 The *Sulfolobus* Tat pathway

1.5.1 Introduction to *Sulfolobus* spp.

Sulfolobales are hyperthermophilic archaea found in sulphur-rich environments such as solfataric fields, hot waters and mud pots. They require acidic conditions (pH 1.0 to 5.0) and have an optimal growth at temperatures between 60 and 90°C [96]. Four genomes of the *Sulfolobales* have been published, *Sulfolobus solfataricus*, *Sulfolobus tokodaii, Sulfolobus acidocaldarius* and *Metallospharea sedula* [97-100]. The *Sulfolobus islandicus* genome is also available although it is still incomplete (http://genome.jgi-psf.org/draft_microbes/suly2/suly2.info.html.).

This thesis specifically looks at *S. tokodaii* and *S. solfataricus. S. solfataricus* was isolated from the acidic hot spring found in Agnano, Napoli, Italy and optimal growth is between 80 and 85ºC at pH 2-4 [101, 102]. While *S. tokodaii* was isolated from the Beppu Hotspings in Kyushu, Japan and optimal growth is slightly lower at 75-80ºC and pH 2.0-3.0 [98, 103]. Both species are obligate aerobes that grow either lithoautotrophically by oxidizing sulphur or chemo-hetrotrophically using reduced carbon compounds [96, 104].

The membrane of thermophilic archaea often consists of a mixture of archaeal ether lipids and so-called tetraethers. Tetraethers are made up of two headgroups fused together by a C_{40} core allowing them to span the entire membrane (Figure 1.11). This allows the membrane to be much more stable at these high temperatures [105]. S*ulfolobus* spp. are unique because they are the only known archaea with a membrane that consists solely of tetraethers, forming a very stable monolayer [106].

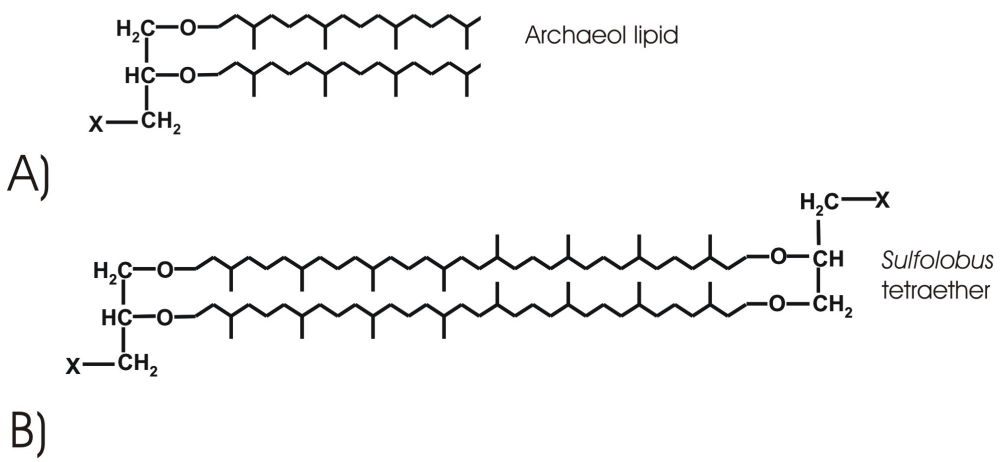

Figure 1.11 Chemical representation of the Archaeal membrane lipids and the Archaeal membrane tetraethers. A) the basic lipid found in the Archaeal membrane. B) tetraethers which are made from two archaeal lipids fused by a C_{40} core. They are bipolar with the X representing a phosphate group on one side (phosphatidylmyoinositol) and a sugar on the other side (β-glucose or β-D-galactosyl-D-glucose) [107]..

The extremely high temperature and low pH that some *Sulfolobus* spp. survive in, as well as its unique monolayer, make this organism very interesting for investigating the archaeal Tat pathway.

1.5.2 The *Sulfolobus* Tat pathway

S. tokodaii and *S. acidocaldarius* each contain one operon that consists of three *tat* genes (*tatA*$_1$, *tatA*$_2$, *tatC*$_1$; denoted as *tatA*$_1$*A*$_2$*C*$_1$). *S. solfataricus* also contains a three gene operon (*tatA*$_1$*A*$_2$*C*$_1$) but there is a second smaller Tat operon consisting of just two genes (*tatA*$_3$, *tatC*$_2$; denoted as *tatA*$_3$*C*$_2$) (Figure 1.12). The second Tat operon is located 244 Kbps downstream of the first.

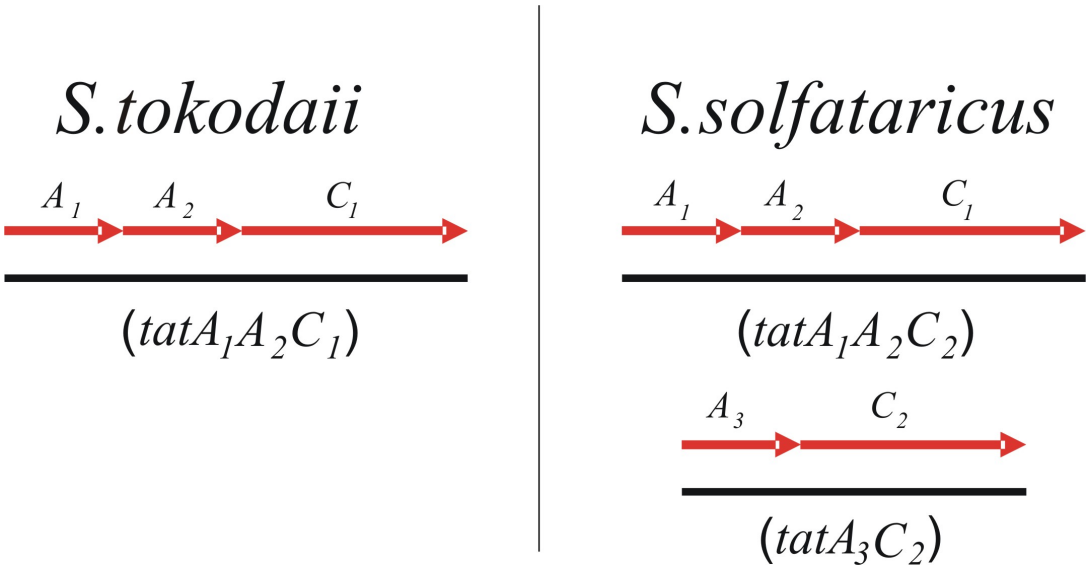

Figure 1.12 The Tat operons found in *S. tokodaii* and *S. solfataricus*. In this thesis, the names above the arrows are used to describe the individual Tat genes. The names in brackets below the operon are used to describe the entire operon.

All three *Sulfolobus* TatC proteins have approximately 80% similarity with each other and 45% similarity with the *E. coli* TatC protein. There is very little similarity between the *Sulfolobus* TatCs and the two TatC proteins found in *Bacillus* (~14%) or *Halobacterium*-NRC1 (~10%). Since TatA and TatB are homologous proteins but have different functions [53], it makes it difficult to determine if they are both present in *Sulfolobus* or if there are two copies of the same *tatA* or *tatB* gene. However, there does appear to be a clear distinction between the *S. tokodaii* and *S. solfataricus* $TatA_1$ and $TatA_3$ proteins compared to the two $TatA_2$ proteins. There is a 65% similarity between $TatA_1$ and $TatA_3$ and there is a phenylalanine prior to the conserved glycine which in Gram negative bacteria is the main distinctive feature of TatA proteins (See Table 1.2). On the other hand, the $TatA_2$ proteins only share 37% similarity with $TatA_1$ and $TatA_3$ and lack this phenylalanine. Notably, the *Sulfolobus* TatA2 proteins also lack the proline residue after that glycine, (which is characteristic for TatB proteins in Gram-negative bacteria). Interestingly, analysis of several other archaeal TatA proteins suggest that along with the conserved phenylalanine there is also a conserved alanine after the glycine that is not seen with bacterial TatA proteins (See Table 1.2).

Organism	Alignment of TatA proteins
S. solfataricus $TatA_1$	--------MLTNPYDW-IIILVVIAVLF**FGA**SKIPELFRS
S. solfataricus $TatA_2$	--------MISNLSDF-LVVVVVFILLMA**G**DKNAGNTTKS
S. solfataricus $TatA_3$	--------MIDNPTDI-IIILVVIAVLF**FG**SSKMPELFRA
S. tokodaii $TatA_1$	-------MALGSVYDA-AIIIVVAIILI**FGA**SKLPEIFRS
S. tokodaii $TatA_2$	--------MLGSLDDF-LIVLLVGILLFA**G**DKNSVNSIKN
E. coli TatA	---------MGGISIWQLLIIAVIVVLL**FG**TKKLGSIGSD
E. coli TatB	---------MFDIGFSELLLVFIIGLVVL**GP**QRLPVAVKT
B. subtilis TatAy	---------MPIGPGS-LAVIAIVALII**FGP**KKLPELGKA
B. subtilis TatAd	--------MFSNIGIPGLILIFVIAIII**FGP**SKLPEIGRA
B. subtilis TatAc	---------MELSFTK-ILVILFVGFLV**FGP**DKLPALGRA
N. pharaonis $TatA_1$	-MIDLLPLQFGIPGGMELLVVVLIAVLL**FGA**DKVPKLART
H. salinarum $TatA_1$	MFTSTPLFIGGLPGGMEMAVVLLIAILL**FGA**NKIPKLARS
H. marismortui $TatA_1$	MLGTSGVLFPGMPGTTEMMVILLIAVLL**FGA**NKIPKLARS

Table 1.2 Alignments of the TatA proteins from a variety of organisms. This table shows the alignments made of the TatA proteins from *E. coli*, *B. Subtilis* [23], *S. solfataricus*, *S. tokodaii*, *N. pharaonis, H. marismortui* and *H. salinarum* . The **F** shows the conserved phenylalanine of the TatA proteins, **G** shows the conserved glycine, **P** shows the conserved proline of the *B. subtilis* TatA and the *E.coli* TatB proteins while the **A** shows the alanine that follows the conserved glycine of the archaeal TatA proteins analysed.

In *S. solfataricus* there are predicted to be 114 secretory proteins which make up 3.9% of the organism's proteome; most of these are probably Sec-dependent [108]. Of these secretory proteins, only six contain a convincing twin arginine motif while in *S. tokodaii* and *S. acidocaldarius* there are only four (Table 1.3). Based on their similarity to proteins from other organisms, all of these are likely to bind to cofactors, which explain why they would have to utilise the Tat system. It is conceivable that the *Sulfolobus* Tat pathway is less stringent than the bacterial system which could cause substrates with weak Tat signal peptides to be missed [109]. However, even if this were to be the case it seems likely that the *Sulfolobu*s Tat pathway only translocates a small portion of secretory proteins.

Protein	Tat signal motif	Name
S. solfataricus		
SSO1580 - Molybdopterin oxidoreductase	LRLN**RRDFLK**ASAAAGLVMGLGYFASKYNFNI FVTRD	Mo1
SSO2795 DMSO reductase (oxidoreductase)	MLKL**SRRDFLK**ISGATAVATAFILGGNSVAKR IFDSVSET	$DMSO_r$
SSO2826 hypothetical protein, (oxidoreductase)	MRLTN**RRDFLK**LLLISSSLLVLGRLSISEMQH AQNGLTPF	Od1
SSO2803_SULSO Rieske iron-sulfur protein-2	DEK**RRKFLK**SLIFGIAAAAVVGVIPGLKVLVP PT	SoxL-2
SSO2660 Rieske iron-sulfur protein-1	MMN**RRTFLR**LYLVVGAAVAIAPLIKPLADYVG YFYNEI	SoxL-1
SSO2971 Quinol oxidase-2, (Rieske iron-sulfur protein-2)	MMN**RRTFLR**LYLVVGAAVAIAPLIKPLADYVG YFYNEI	Qo2
S. tokodaii		
ST0108 Rieske iron-sulfur protein	MK**RRDFIR**LAMIAGGAVAISPLFSPLF NYMGYYYNEL	
ST1010 Sulfite oxidase (homologue of SS02826)	MK**RRDFLK**ALFISTSLLVIGRLSLYEYNNYQH VQNALT	
ST1666 Rieske iron-sulfur protein	K**RRKFLK**GLIFGIAAATVAAIIPGVEVVVPPQ VAAVS	
ST1840 DMSO reductase (homologue of SS02795)	LKL**TRRDFLK**LSGITALTTAILLEGGSVIDKV FRGQEDFT	
S. acidocaldarius		
SA70606526 hypothetical protein	LSPS**SRRKVFI**SCNFLISVLLIGFYLSTPLTV WIFTAIAG	
SA70606666 hypothetical protein	MKTVFN**DRRAIIY**GVVFSIASYVTAAILAVGS SVYALGSFL	
SA70607819 hypothetical protein (Molybdopterin oxidoreductase)	MSQLKL**SRRDFLK**ISGFAGLGTAVFLSTPSVV DKIFRPIAETTSY	
SA70607819 Rieske domain	M**DRRTFLR**LYLLVGAAIAVAPVIKPALDYVGY FYSELGSL	

Table 1.3 The signal peptides of predicted *S. solfataricus, S. tokodaii* and *S. acidocaldarius* Tat proteins. This table shows the signal peptides of the 6 *S. solfataricus,* 4 *S. tokodaii* and 4 *S. acidocaldarius* predicted proteins. The Twin arginines are shown in red and the rest of Tat motif is indicated in bold while the underlined regions indicate the hydrophobic domains. Short hand names for the *S. solfataricus* Tat substrates are also indicated.

1.5.3 The hypothesis of the *Sulfolobus* Tat Pathway

The putative Tat substrates in *S. solfataricus* and *S. tokodaii* are mainly Rieske proteins that bind 2Fe-2S iron sulphur clusters and molybdopterin oxidoreductases such as DMSO reductase that bind molybdenum [110, 111] Therefore, it seems likely that in *Sulfolobus* the Tat pathway is specifically involved in the translocation of fully folded proteins containing complex cofactors. Since Rieske iron-sulphur proteins are

involved in energy metabolism and oxidoreductases are often involved in respiration, it is quite possible that the Tat pathway plays an important role in *Sulfolobus* and its absence may be lethal to the cell or at least affect its phenotype [112].

It is not possible to determine if the pathway resembles the *E. coli* or *B. subtilis* system, or whether it has developed it own unique method. The fact that there are two Tat operons in *S. solfataricus* indicates that there may be some substrate specificity as found in *B. subtilis.* However, as mentioned previously, alignments indicate that there is more similarity with the *E. coli* TatC protein and also the possible presence of a TatB protein. It is also conceivable that because of the vast differences in conditions and the unique chemical composition of the membrane that there may also be novel proteins and features present.

1.6 The Halophilic Tat Pathway

1.6.1 Introduction to *Halophilic* archaea

Haloarchaea are salt-loving organisms that require more than 15% NaCl to grow and are often found in environments such as salt lakes and salt evaporation pools [113]. They are heterotrophic and normally respire by aerobic means. The high external salt concentrations make it vital for the cell to maintain osmotic balance. This is done by accumulating K^+ in the cytoplasm to match the Na^+ in the environment using a process known as the "Salt-in strategy" [114, 115]. This has meant that halophilic proteins and cellular processes have adapted to function in these extremely saline conditions. As a result, most Haloarchaea are unable to survive outside of their native high-salt environments [116].

1.6.2 The Halophilic Tat pathway

In *Halobacterium* sp. NRC-1, it was found that 65-70% of the secreted proteins contain a predicted twin arginine motif [117, 118]. Since this pathway is the major route of secretion it is of no surprise that it has been found to be essential in haloarchaea [94, 119]. It seems likely that the Tat pathway is used extensively due to the high amounts of KCl found in the cytoplasm due to the "Salt-in strategy". This has meant that halophilic proteins have evolved to contain high amounts of acidic residues but only a few basic residues [120, 121]. Many of these acidic residues are located on the surface of halophilic proteins, and this negative surface charge is probably important in the binding of essential water molecules and prevention of aggregation through electrostatic repulsion [121, 122]. High salt concentrations have a "salting-out effect" and it is therefore conceivable that halophilic proteins fold very rapidly in the cytoplasm (although is has to be noted that there are no reports in the literature on the rate of folding of halophilic proteins). If so, this would mean that many secretory proteins in haloarchaea will fold before translocation and that in turn would explain the extensive use of the Tat system for export in halophilic archaea. It has also been suggested that halophilic proteins require the assistance of cytoplasmic chaperones due to the high KCl concentrations and therefore have to fold before translocating [123],

but it has to be noted that a number of haloarchaeal proteins are able to fold spontaneously *in vitro* such as the halophilic malate dehydrogenease (HmMalDH) from the haloarchaeon *Haloarcula marismortui* [123] and AmyH from *Haloarcula hispanica* (See Chapter 6).

All haloarchaea investigated contain one *tatA* and two *tatC* genes but no *tatB* gene has been found [124]. *Haloferax volcanii* contains a second TatA gene (denoted *tatAo*) but it is not required for translocation to occur [125]. The two *tatC* genes are adjacent to each other but are transcribed by two diverging promoters (Figure 1.13) indicating that they form a functional unit [119]. Unlike in other organisms, *tatA* is not present in the same operon but is found elsewhere on the chromosome. *Halobacterium* sp. NRC-1 TatA protein shares approximately 36% similarity (19% identity) with the *E. coli* and *Bacillus* TatA proteins while there is only approximately 28% similarity (13% identity) for $TatC_1$and 14% similarity (7% identity) for $TatC_2$ when compared to these bacterial TatC's.

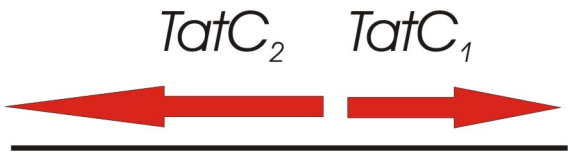

Figure 1.13 The Halophilic Tat operon. The Tat operon found in *Halobacterium salinarum*, *Haloarcula marismortui*, *Haloferax volcanii*, and *Natromonas pharaonis*. *$TatC_2$* encodes the larger TatC protein although the names are reversed for *H. marismortui*. TatA is located 107 kb upstream of the TatC genes.

There are two main differences between the haloarchaeal and the bacterial TatC proteins. $TatC_1$ contains an uncharacteristically long N-terminal cytoplasmic domain and $TatC_2$ consists of two TatC domains joined by a linker domain (Figure 1.14) [21]. The role of $TatC_2$ is not clear, but it is noteworthy that this type of TatC protein is only found in haloarchaea, and it is therefore possible that this is an adaptation to the extremely high salt concentrations. It is conceivable that these adapted TatC proteins allow for the increased use of this pathway, for example by improving the speed of formation/dissociation of the TatAC or TatC-substrate complexes or by allowing the system to recognise or secrete proteins that are in folding states not accepted by other organisms [94, 125].

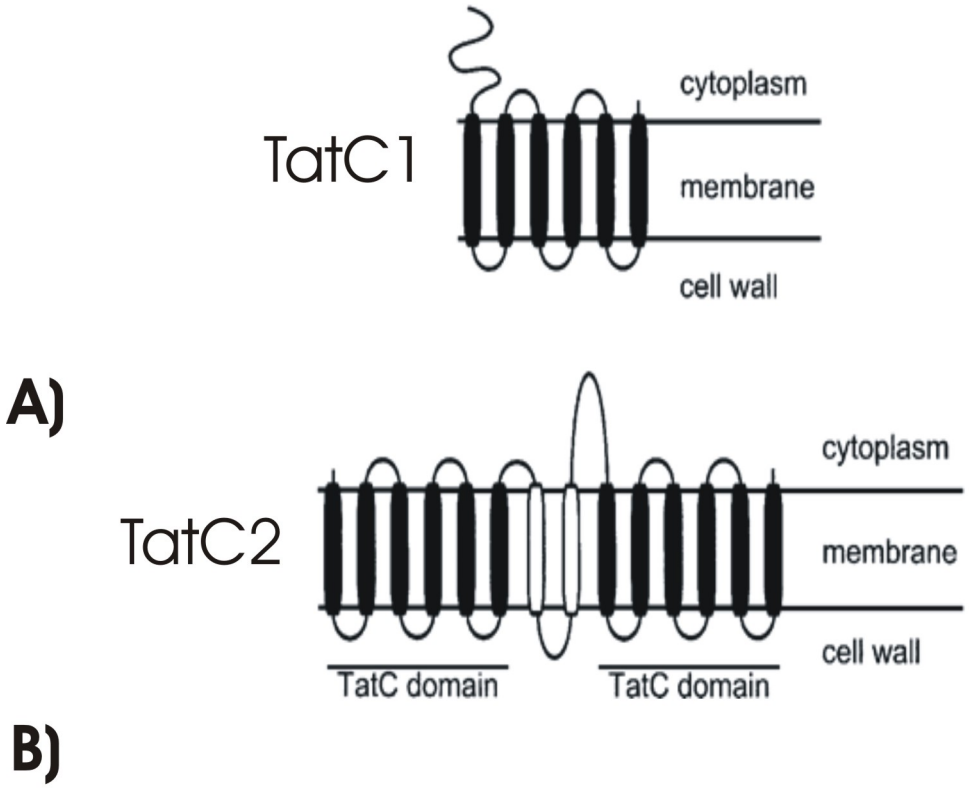

Figure 1.14 The predicted formation of $TatC_1$ and $TatC_2$ in the Halophilic membrane. This diagram shows the predicted membrane topologies of $TatC_1$ and $TatC_2$. Note the large N-terminal cytoplasmic domain of $TatC_1$ and the linker domain of $TatC_2$

This Thesis specifically looks at the Tat pathway of the haloarchaeon *H. hispanica,* which was isolated from solar salterns near Alicante, Spain. The organism is chemo-organotrophic and grows between 25 and 50 ºC (Optimum 35-40 ºC) and pH 6.0 to 8.0 (Optimum 7.0). The optimum salt concentration for growth is at 25% (w/v) although they are capable of growing between 15% (w/v) and salt saturation [126]. Unlike most Archaea, haloarchaea are amenable to genetic manipulation as there are several genetic tools available to work with for example transformation procedures, shuttle vectors and knockout strategies [127, 128].

1.6.3 The hypothesis of the Halophile Tat Pathway

The high concentration of salt found in the haloarchaeal cytoplasm has meant that many cellular processes have had to adapt to continue to function. This is true for the Tat pathway which has become the major route of translocation, this is a huge

contrast to what is known about other Tat pathways. It is therefore likely that the system has evolved to allow it to cope with the increased number of Tat substrates. For example, the twin arginine motif may be less stringent, or the prerequisite of being fully folded may be less strict to allow more proteins to utilise the pathway.

1.7 The aims of this thesis

Archaea have a particular scientific interest because of their ability to survive in such extreme environments that were previously thought to be inhabitable. Since protein transport is an essential and important cellular process, this thesis aims to investigate the Tat pathway of the Archaeal species, *Sulfolobus* and *Haloarcula*. It is hoped that by doing so it will give an indication into the extent that these organisms have had to adapt to enable them to survive in these extremely high temperatures or high salt conditions.

Chapter 2: Materials and Methods

2.1 Plasmids

Plasmid	Details	Reference
pBAD24	*E. coli* cloning vector *araC, Amp^r*	Guzman *et al*., (1995) [129]
pSA1A2C1	pBAD24, *S.so tatA$_1$A$_2$C$_1$*	This Study
pSA3C2	pBAD24, *S.so tatA$_3$C$_2$*	This Study
pTA1A2C1	pBAD24, *S.to tatA$_1$A$_2$C$_1$*	This Study
pA1	pBAD24, *S.so tatA$_1$*	This Study
pA2	pBAD24, *S.so tatA$_2$*	This Study
pA3	pBAD24, *S.so tatA$_3$*	This Study
pC1	pBAD24, *S.so tatC$_1$*	This Study
pC2	pBAD24, *S.so tatC$_2$*	This Study
pAYC-RIL	pAYC plasmid encoded rare arginine, isoleucine and leucine codons	Stratagene
pSPTA1	pSY1 (*E. coli-H.vol* shuttle plasmid), *SptA*	Yang *et al*., (2003) [130]
pSPTAHIS	pSY1, *SptA* with His Tag	This Study
pAMY	pSY1, *amyH*	This Study.
pAMYKK	pSY1, *amyH* KK mutant	This Study.
pAMYKR	pSY1, *amyH* KR mutant	This Study.
pAMYRK	pSY1, *amyH* RK mutant	This Study.
pAMYAS1	pSY1, *amyH* active site 1 mutant	This Study.
pMZ1	*E. coli* entry vector to pMJO5, *Amp^r,* a*raC*	Albers *et al*., (2006) [131]
pMA1A2C1	pMZ1, *S.so tatA$_1$A$_2$C$_1$*	This Study
pMA3C2	pMZ1, *S.so tatA$_3$C$_2$*	This Study
pMJO5-03	*E. coli* - *S.so* shuttle vector	Albers *et al*., (2006) [131]
pEXPA1A2C1	pMJO5, *S.so tatA$_1$A$_2$C$_1$*	This Study
pEXPA3C1	pMJO5, *S.so tatA3C2*	This Study
pET2268	*E. coli* - *S.so* shuttle vector, *Amp^r, lacS*	Albers and Driessen (2007) [132]
pKOA1A2C1	pET2268, *S.So tatA$_1$A$_2$C$_1$* up/dn flanks	This Study
pKOA3C2	pET2268, *S.So tatA$_3$C$_2$* up/dn flanks	This Study
pET21a	*E. coli* expression vector containing a T7 promoter	Webb et al., (1999) [133]
pET-AMY	Pet21a, *AmyH* gene behind the T7 promoter	This Study
pET-AMYKK	Pet21a, *AmyH* KK mutant gene behind T7 promoter	This Study

Table 2.1 The plasmids used during this project. The details of the plasmids that were used during this project. Also mentioned is if the plasmid was constructed or if it was kindly provided for use in this study.

All of the plasmid constructs that were used in this study are shown in Table 2.1

2.2 Bacterial and Archaeal Strains

2.2.1 Bacterial strains

All of the bacterial strains that were used in this study are shown in Table 2.2

Strain	Genotype	Reference
E. coli MC4100	F,Δ*lacU169*,araD139,rpsL150,relA1,ptsF,rbs, flbBS301	Casadaban (1976) [134]
E. coli K12 JM109	F′ *traD36 proA*$^{+}$*B*$^{+}$ *lacI*^{q} Δ*(lacZ)M15/* Δ*(lac-proAB) glnV44 e14*$^{-}$ *gyrA96 recA1 relA1 endA1 thi hsdR17*	Yanisch-Prerron *et al*., (1985) [135]
E. coli ElectroMax stble4 Cells	*mcr*A Δ(*mcr*BC-*hsd*RMS-*mrr*) *rec*A1 *end*A1 *gyr*A96 *gal*$^{-}$ *thi*-1 *sup*E44 λ^{-} *rel*A1 Δ(*lac-pro*AB)/F'*pro*AB^{+} *lac*I^{q}ZΔM15 Tn*10* Tet^{R}	Dower *et al*., (1988) [136]
E. coli K12 ER2508	F^{-} *ara-14 leuB6 fhuA2* Δ*(argF-lac)U169 lacY1 lon::miniTn*10(Tet^{R}) *glnV44 galK2 rpsL20*(Str^{R}) *xyl-5 mtl-5* Δ*(malB) zjc::Tn*5(Kan^{R}) Δ*(mcrC-mrr)*$_{HB101}$	Kowit and Goldberg (1977) [137]
E. coli K12 ER2925	*ara-14 leuB6 fhuA31 lacY1 tsx78 glnV44 galK2 galT22 mcrA dcm-6 hisG4 rfbD1 R(zgb210::Tn10)TetS endA1 rpsL136 dam13::Tn9 xylA-5 mtl-1 thi-1 mcrB1 hsdR2*	Palmer and Marinus (1994) [138]
E. coli DADE	MC4100; Δ*tatABCDE, Ara*^{r}	Wexler *et al*., (2000) [46]
E. coli XL10 Gold	Tet^{r} Δ*(mcrA)183* Δ*(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F' *proAB lacI*^{q}*ZΔM15 Tn10* (Tet^{r}) Amy Cam^{r}]^{a}	Trauger and Walsh (2000) [139]

Table 2.2 The *E. coli* strains used during this project. All of the strains that were used during this project and also information about their genomes.

2.2.2 Haloarchaeal strains

All of the haloarchaeal strains that were used in this study are shown in Table 2.3

Strain	Information	Reference
Haloarcula hispanica	*H. hispanica* DSM No.4426	Juez *et al*., (1986) [126]
Haloarcula hispanica B3	An over expressing AmyH mutant strain of *H. hispanica* DSM No.4426	Hutcheon *et al*., (2005) [113]
Haloferax volcanii H26	Δ*pyrE2* mutant of *H. volcanii* DS70	Allers *et al*., (2004) [140]

Table 2.3 The halophile strains used in this project. This table shows all of the strains of *Haloarcula* and *Haloferax* that were used during this project.

2.2.3 *Sulfolobus* strains

All of the *Sulfolobus* strains that were used in this study are shown in Table 2.4

Strain	Information	Reference
Sulfolobus solfataricus P2	Wild type	She *et al*., (2001) [99]
Sulfolobus solfataricus M16	*S. solfataricus* P2, Δ*pyrEF*	Albers *et al*., (2006) [141]
Sulfolobus solfataricus PBL2025	*S. solfataricus* P2, Δ*lacS*	Worthington *et al*., (2003) [142]
Sulfolobus tokodaii	Wild type	Kawarabavasi *et al*., (2001) [98]

Table 2.4 The sulfolobus strains used during this project. This table shows all of the Sulfolobus strains that were used during this project.

2.2.4 Alkaliphile strains

All of the alkaliphile strains that were used in this study are shown in Table 2.5

Strain	Information	Reference
Natronomonas pharaonis	*N. pharaonis* WDG53	Kindly provided by Dr. W.D.Grant [143]
Natronococcus occultus	*N. occultus* WDG55	Kindly provided by Dr. W.D.Grant [143]

Table 2.5 The alkaliphile strains used during this project. This table shows all of the alkaliphile strains that were used during this project.

2.3 Chemicals

All chemicals were purchased from Fisher Scientific or Sigma unless stated otherwise. The suppliers for enzymes, kits and equipment are mentioned in the relevant sections of this Chapter.

2.4 Cell culture methods

2.4.1 *E. coli* growth conditions

All *E .coli* strains were grown on Luria-Bertani Medium (LB) as described by Sambrook and Russell (1989) [144]. If required, 100 µg/ml of ampicillin or 10 mM arabinose were added. All cultures were grown overnight at 37°C apart from *E. coli* ElectroMax stble4 cells which were grown following manufacturer's instructions.

2.4.2 Haloarchaeal growth conditions

Unless stated, all *H. volcanii* cultures were grown at 45°C in 18% HV-YPC media as described by Allers *et al*., (2004) [140]. *H. hispanica* requires higher salt concentrations so were grown on 23% salt HV-YPC media. If required, novobiocin was added at 0.3 µg/ml. For translocation assays, 1% skim milk (SptA translocation assay) or 0.5% starch (AmyH translocation assay) were also added. After 2 days incubation for the AmyH translocation assay, plates were stained for 30 seconds with iodine solution (2% Potassium iodine, 0.2% Iodine).

2.4.3 *Sulfolobus* growth conditions

Unless stated, all *Sulfolobus* strains were grown at 75°C on 0.1% tryptone, 0.4% sucrose Brock salt basal medium, pH 3 (Brock media) [145]. Prior to transformation, the medium for Sulfolobus M16 supplemented with 10 µg/ml uracil.

2.4.4 Alkaliphile growth conditions

Natronomonas pharaonis and *Natronococcus occultus* cultures were grown at 45°C on media containing 1% w/v yeast extract, 0.75% w/v casamino acids, 0.3% w/v $Na_3C_6H_5O_7 \cdot 2H_2O$, 0.2% w/v KCl, 0.1% w/v $MgSO_4.7H_20$, 20% w/v NaCl, 1.85% w/v Na_2CO_3, 0.00036% w/v $MnCl_2.4H_20$ and 0.005% w/v $C_6H_8O_7$ Fe NH_3 pH 9-9.5. For growth on plates then 2% agar was added. Cultures were generally grown for 3-5 days.

2.4.5 Viable count

Viable counts were performed so that optical density could be related to colony forming units (cfu) of *Haloferax* and *Haloarcula*. Cultures were grown to an OD_{600} of 1 and duplicate plated onto Hv-YPC agar at 10^{-4}, 10^{-5} and 10^{-6} dilutions. Using the dilution that fell within the countable range (20-100) it was possible to approximate the colony forming units at a specific absorbance.

2.4.6 Harvesting cells

Unless stated otherwise, cells were separated from the media by centrifugation at 6 000 g

2.5 Nucleic Acid purification

2.5.1 Small scale plasmid purification

The viral vector pMJO5 was isolated using alkaline lysis and phenol/chloroform purification described by Sambrook and Russell (1989) [144] . All other plasmids were isolated using the Nucleospin Plasmid Kit (Macherey-Nagel) following manufacturer's instructions.

2.5.2 Agarose gel electrophoresis and gel purification

Agarose gel electrophoresis was carried out as described by Sambrook and Russell (1989) [144]. 0.9% agarose gels containing 0.5 µg/ml ethidium bromide were run at 120V and DNA was visualised under UV light. DNA gel purification was performed using the Nucleospin Extract II Kit (Macherey-Nagel) following manufacturer's instructions.

2.5.3 *Sulfolobus* chromosomal DNA purification

Cell lysis followed the protocol by Yeats *et al*., (1982) [146] but 0.15% Triton X-100 was included with the 1.6% sodium lauryl sarcosine. Once cells were lysed, plasmid DNA was purified using phenol/chloroform extraction. For this, one volume of phenol/chloroform/isoamyl-alcohol (25:24:1) was added to the sample and then centrifuged for 5 minutes. The upper aqueous phase was removed to a fresh tube and the phenol/chloroform extraction process was repeated for a second time.

The DNA was then precipitated by the addition of $1/10^{th}$ the volume of 3 M sodium acetate pH 5.2 and 2.5 volumes of ethanol. The sample was left on ice for 30 minutes and centrifuged at 14 000 g (4ºC) for 20 minutes. Finally the pellet was washed with 500 µl of 70% ethanol, spun down and air dried. The DNA was then re-suspended in 10-20 µl H_2O. Finally, the DNA was re-suspended in TE buffer (10 mM Tris/HCl, 1 mM EDTA pH 8) containing 10 µg/ml of RNase.

2.5.4 DNA quantification

DNA was quantified by measuring the absorption at 260 nm using the Eppendorf Bio-Photometer. An absorbance of 1.0 indicated 50 µg/ml of double stranded DNA.

2.5.5 Automated DNA sequencing

Samples to be sent for sequencing contained 150 ng/µl of plasmid DNA and 15 pmols of primer in a total volume of 15 µl. Sequencing was performed by Eurofins MWG Operon, UK.

2.6 Cloning techniques

2.6.1 Polymerase chain reaction (PCR)

Typically PCR involved a 50 µl reaction using Platinum *Taq* DNA Polymerase High Fidelity kit (Invitrogen) and stocks of dNTPs (Fermentas). Reactions were set up following Invitrogen's instructions except for amplifying *Sulfolobus* A-T rich DNA where the concentration of buffer was doubled. For amplifying G-C rich *Haloferax* or *Haloarcula* DNA, DyNAzyme™ EXT DNA Polymerase (New England Biolabs) was used and also followed manufacturer's instructions. PCR was performed in a Eppendorf Mastercycler gradient machine and followed Invitrogens PCR amplification protocol. The annealing temperatures varied depending on primers but generally a temperature of 60°C was sufficient. All oligonucleotides were synthesised by Invitrogen and are noted in the relevant chapters.

2.6.2 Site directed mutagenesis

Site directed mutagenesis was performed using Stratagene's Quick Change II XL Site-Directed Mutagenesis Kit. Primers were designed using Stratagene's Quick Change primer design programme. To prevent primer self annealing, two-stage site directed mutagenesis was performed. Two separate 50 µl PCR reactions were made following manufacturer's instructions but each sample only contained one of the complimentary primers and 0.5 µl of *Pfu* Ultra HF DNA polymerase. These samples went through 3 cycles of the recommended amplification programme. 25 µl of each reaction was then pooled together and 1 µl of *Pfu* Ultra HF DNA polymerase was added. The rest of the experiment then followed Stratagene's protocol.

2.6.3 Restriction enzyme digestion

All restriction enzymes were purchased from Invitrogen or Fermentas, and reactions were carried out according to manufacturer's instructions. DNA to be ligated was always digested overnight. After digestion, DNA was purified using the Nucleospin

Plasmid Kit (Macherey-Nagel) or from gel using the Nucleospin Extract II Kit (Macherey-Nagel) following manufacturer's instructions.

2.6.4 Ligation of DNA into cut plasmid vector

Ligations were performed using Invitrogen's T4 DNA ligase and reactions were set up following manufactures instructions. 10-15 ng of digested vector and 30-45 ng of digested insert were always used. Reactions were incubated overnight at room temperature.

2.7 Transformation of competent cells with plasmid DNA

2.7.1 Transformations of Calcium competent *E. coli*

E. coli strains JM109, DADE, MC4100 and ER2508 were made competent by treatment with $CaCl_2$ as described by Sambrook and Russell (1989) [144]. Either 5 µl of ligation mixture or 2 µl of plasmid DNA were mixed with 50 µl of competent cells. Cells were left on ice for 30 minutes and then heat-shocked at 42ºC for 30 seconds. Samples were then returned to ice for a further 2 minutes. 1 ml of SOC medium (Sambrook and Russell,1989) [144] was then added to the samples and incubated for 1 hour at 37ºC. Finally, 10 µl, 100 µl and 200 µl of these transformed cells were plated onto LB plates containing the appropriate antibiotic and incubated overnight at 37ºC. Commercial *E. coli* XL10 Gold (Stratagene) was transformed according to manufacturer's instructions.

2.7.2 Transformations of Electro competent *E. coli*

E. coli ElectroMAX Stble4 Cells (Invitrogen) were transformed according to manufacturer's instructions using an electroporator (Eppendorf 2510).

2.7.3 *S. solfataricus* transformation

Sulfolobus strains were made electro-competent following the protocol by Schelert *et al*., (2004) [147]. 100-300 ng of DNA was then added to 50 µl of competent cells and electroporated at 1.5 kV, 25 µF and 400 Ω, the time constant was always around 10 ms. *S. solfataricus* M16 was transformed when attempting protein expression. Cells were then allowed to regenerate at 75°C for 1 hour in 1 ml of 0.1% Tryptone Brock media containing 10 µg of uracil. After the hour, Cells were transferred to 50 ml of media and allowed to grow for a further 2 days. *S. solfataricus* PBL2025 cells were used for creating knockout mutants. After transformation, cells were added to 1 ml of water and left for 10 minutes at 75°C. They were then transferred to 50 ml of 0.4% lactose Brock media and allowed to grow for a further 12 days. Cells were then plated on 0.4% lactose plates and incubated at 75°C for 6 days. Successful transformants were then confirmed by spraying with 5 mg/ml X-gal for blue colony selection

2.7.4 *H. volcanii* transformation

Transformation of *H. volcanii* followed the protocol of Cline *et al*., (1989) [148]. Transformed cells were then plated onto selective novobiocin Hv-YPC plates and incubated at 42ºC for 5 days.

2.8 RNA Techniques

2.8.1 RNA extraction from *S. solfataricus* and RT-PCR

Sulfolobus cultures were grown to an OD_{660} of 0.5 in 0.4% tryptone or 1% glucose Brock media. RNA was then purified using the RNeasy mini kit (Qiagen) following manufacturer's instructions. The samples were treated with 1 unit of DNase I (Roche, RNase free) per µg of RNA and incubated at 37ºC for 30 mins. The DNase I enzyme was then inactivated at 75ºC for 5 minutes.

The Expand RT-PCR kit (Roche) was then used to perform RT-PCR on 0.5 µg of RNA following manufacturer's instructions. The resulting cDNA was amplified using *Taq* DNA polymerase as described in Section 2.6.1. The primers used for RT-PCR on *S. solfataricus* mRNA are shown in Table 2.6.

RT-PCR	Forward Primer (Only used for PCR)	Reverse Primer
$DMSO_r$	ATAGTATACACTAATTGTTTCCAATGCTTA	ATTGCCCATTACTGCTTCTAGCC
MO1	ATGAAGAGTTACAACCGGGTT	AACTTTTTGTGCTAGGAATTGGAA
OD1	TAATATCCTCATCGTTATTGGTCTT	TAGCCTAGAATGTTTTTAGATGT
QO2	AGTTGCAATAGCACCCTTAATAA	ATCAGCTGGAGTAGTTCTGG
SoxL-1	AGGTATAGGAGCCGCAGCA	CATCATATGAAGATAGATACGCATACCCTG
SoxL-2	TACCAGGACTTAAAGTATTGGTC	GCTAACGGAGCTACCGAA
$A_1A_2C_1$	ATATGAACATGAGAAGAAACTTTTCAAG	ATGCTTCTCAAATTGTGGATGAC
A_3C_2	TATTTGATCCAATATATGCGTCGC	TGGATGACTCCATCCGGATC

Table 2.6 Primers used to perform RT-PCR on the *S. solfataricus* $A_1A_2C_1$, A_3C_2 and Tat substrate mRNA. All primers are annotated 5'-3'.

2.8.2 RNA extraction from *E. coli* and RT-PCR

The *E. coli* culture was grown to an OD_{660} of 0.7 and mRNA was purified using the RNeasy mini kit (Qiagen) following manufacturer's instructions. RT-PCR using the M-MLV reverse transcriptase (Invitrogen) was then performed on 0.5 µg of mRNA, again following manufacturer's instructions. Finally the resulting cDNA was amplified using the *Taq* DNA polymerase as described in Section 2.6.1. Primers used in this project for RT-PCR on *E. coli* are shown in Table 2.7.

RT-PCR	Forward Primer (Only used for PCR)	Reverse Primer	Details
$A_1A_2C_1$	TAGAGTAGAGGCCGAAATGG	AGGCAGAACCTAAGGCAAAC	To detect *S. solfataricus* $A_1A_2C_1$ mRNA
A_3C_2	GGAGTTCAAGAAGGGACAACTAGAAG	CGAGGTCAATAGTGCGAGGTAAAGTG	To detect *S. solfataricus* A_3C_2 mRNA

Table 2.7 Primers used to perform RT-PCR of the *S. solfataricus* $A_1A_2C_1$ and A_3C_2 mRNA expressed in *E. coli*. All primers are annotated 5'-3'.

2.9 Protein gel techniques

2.9.1 SDS-polyacrylamide gel electrophoresis (PAGE)

Protein samples were separated by SDS PAGE as described by Laemmli *et* al., (1970) [149]. Samples taken straight from halophilic cultures were diluted two-fold with loading buffer before loading to reduce the salt concentration. The EZrun Pre-stained Rec protein ladder (Fisher) was always loaded so that protein size could be determined. 10% acrylamide gels were used for all AmyH work while 12.5% acrylamide gels were used for *Sulfolobus* work. Native gels were made in the same way as SDS page gels except that they lack the SDS.

2.9.2 Western Blotting

After proteins were resolved by SDS PAGE, they were transferred onto a polyvinylidene difluoride membrane (PVDF; Millipore) using a semi-dry western blotting system. The blotter was set up with a stack of 2 layers of Whatman chromatography paper (Schleicher&Schuell), then the PVDF membrane activated in methanol, the protein gel and then finally 2 more layers of Whatman paper. The PVDF membrane and Whatman paper had previously been soaked in Towbin buffer (25 mM Tris, 192 mM Glycine, 20% Methanol). The blotter was run at 20 V, 150 mA for 2 hours.

The membrane was blocked with 5% skim milk (Marvel) dissolved in TBST (150 mM Nacl, 20 mM KCl, 25 mM Tris pH 7.4 and 0.05% Tween-20/litre) for two hours. The primary antibody (Table 2.8) was then added at the appropriate concentration and the blot was left shaking overnight at room temperature. The next day the membrane was washed 3 times in TBST for 5 mins and once for 15mins. The secondary antibody (Table 2.8) was then added at the manufacturer's recommended concentration and left shaking for a further 2 hours. Finally the blot was washed with TBST and signal detection was performed using the Supersignal West Pico-chemiluminescent Substrate Kit (Pierce) following manufacturer's instructions.

Primary antibody	Secondary antibody
Anti-His	Anti-Mouse
Anti-Strep	Anti-Rabbit
Anti-AmyH	Anti-Rabbit

Table 2.8 The antibodies that were used during this project. Shown in the table are the primary antibodies and their corresponding secondary antibodies. Anti-His, Anti-Mouse and Anti-Rabbit were from Promega, Anti-Strep was from IBA BioTAGnology and Anti-AmyH as described by Hutcheon *et al*., (2005) [150].

2.9.3 Southern Blotting

Chromosomal DNA was digested with the desired restriction enzymes and separated by agarose gel electrophoresis. The DNA was blotted onto a Hybond-N^+ membrane (Amersham Pharmacia Biotech) using alkaline transfer as described by Sambrook and Russel (1989) [144]. The desired DNA probes were amplified by PCR and were labelled and hybridised to the blot using the Amersham AlkPhos Direct labelling reagent kit (GE Healthcare) following manufacturer's instructions. Signal detection used the Amersham CDP-Star detection reagent (GE Healthcare) and again followed manufacturer's instructions. Specific restriction enzymes and probes used in this project are noted in the relevant chapters.

2.9.4 Coomassie staining

The Protein sample was separated by SDS PAGE and then fixed for 30 minutes at room temperature in 40% methanol, 10% acetic acid. It was then transferred into Coomassie stain (10% acetic acid, 0.025% Coomassie G-250) and left shaking for 1 hour. Finally the gel was destained overnight (20% methanol, 7% acetic acid) and dried in cellophane wrap.

2.9.5 TMAO Reductase-methyl Viologen Linked gel based assay

The *E. coli* culture was fractionated (Section 2.10.1) and then separated on a 10% native polyacrylamide gel. Immediately after electrophoresis the gel was placed into 100 ml of nitrogen-saturated phosphate buffer pH6.5 (70 mM sodium diphosphate, 30 mM disodium phosphate) containing 1 ml of methyl viologen. The gel was then stained with 10 ml of sodium dithinonite and left for 30 minutes. Finally, the gel was transferred into 100 ml of phosphate buffer pH6.5 containing 2 ml of TMAO. As soon as white bands appeared the gel was photographed as the reaction continues until the whole gel turns white.

2.9.6 Fluorography with 2,5-diphenyloxazole (PPO)

Protein samples containing [^{35}S]-methionine were run on a 10% acrylamide SDS gel and then fixed in Solution A (30% methanol, 10% acetic acid) for 30 minutes. The gel was then soaked in glacial acetic acid for 5 minutes before being transferred to Solution B (20% 2,5-diphenyloxazole in acetic acid) and incubated for 1.5 hours. Finally the gel was soaked in water for 30 minutes and then dried on Whatman paper for 2 hours under vacuum at 80°C. For signal detection the gel was exposed to an x-ray film at -80°C and developed the next day.

2.10 Protein expression and purification

2.10.1 *E. coli* cell lysis and fractionation

E. coli cells were grown until an OD600 of 0.8 and then collected by centrifugation at 8 000 g. They were then spheroplasted using the EDTA/lysozyme/cold osmotic shock method described by Randall and Hardy (1986) [151]. Spheroplasts were collected by centrifugation for 1 minute and the resulting supernatant was the periplasmic fraction. The cellular pellet was re-suspended in Buffer (50 mM Tris-acetate pH 8.2 and 2.5 mM EDTA) and sonicated for 15 minutes using a Branson Sonifier 250 set to a duty cycle of 40% and an output of 4. Centrifugation was performed for 10 minutes at 8 000 g to remove the cellular debris. The cytoplasmic and membrane fractions were then separated by centrifuging the supernatant at 70 000 g for 30 minutes. Finally, the membrane pellet was solubilised overnight at 4ºC in Buffer (50 mM Tris pH 7.5, 500 mM NaCl and 1% n-Dodecyl-β-D-maltoside ULTROL Grade (DDM)).

A protease inhibitor cocktail (Complete – EDTA, Roche Molecular Biochemicals) was included in all buffers after cell lysis to minimise protein degradation.

2.10.2 Haloarchaea cell lysis and fractionation

Cells were harvested by centrifugation at 8 000 g and re-suspended in Buffer (50 mM Tris pH 7.5, 1 M NaCl, 1 mM EDTA) containing a protease inhibitor cocktail (Complete, Roche Molecular Biochemicals). The samples were then sonicated continuously for 15 seconds and cell debris removed by centrifugation at 8 000 g for 10 minutes. The membrane and cytoplasmic fractions were separated by centrifugation at 70 000 g for 30 minutes using the Optima TLX 120 000 Ultracentrifuge (Beckman Coulter). Finally, the membrane was re-suspended in Buffer containing 0.5% Triton X-100.

2.10.3 Purification of His-tagged proteins on a metal chelating column

E. coli cultures containing the desired expression construct were grown at 37ºC for 3 hours. Arabinose was then added and cells were grown to an OD_{600} of 1. Cells were then lysed as described in Section 2.10.1. The HisPur Cobalt resin (Pierce) was used for purification and followed manufacturer's instructions. The only difference was that all buffers contained a protease inhibitor cocktail (Complete – EDTA, Roche Molecular Biochemicals) and 0.1% digitonin (Calbiochem) or 1% DDM (Calbiochem).

2.10.4 Purification using a β-cyclodextrin-sepharose column

AmyH was purified using a β-cyclodextrin-sepharose column as described by Hutcheon *et al*., (2005) [150]

2.10.5 Gel filtration

Gel filtration was performed using the FPLC Chromatography system (GE healthcare) using a Superose 6 column.

2.10.6 Protein dialysis

For samples less then 3 ml, dialysis was performed overnight using the D-tube Dialyzer Midi tubes (3.5 kDa MWCO,Pierce). For larger volumes, dialysis was performed overnight using the Snakeskin Pleated Dialysis tubing (10 kDa MWCO, Pierce). Both methods were performed at 4ºC following manufacturer's instructions.

2.10.7 Concentrating protein samples

The Vivaspin 15R column (5 kDa MWCO, Sartoruis) was used to concentrate volumes of protein less then 50 ml. For the 2 litre samples the Vivaflow 200 tangential filter unit (VivaScience) was used. Both methods followed manufacturer's instructions.

2.10.8 Determining protein concentration

Protein concentration was determined using the BCA protein assay reagents (Pierce) following manufacturer instructions. Known bovine serum albumin (BSA) concentrations were also measured as a reference.

2.10.9 TCA precipitation

100% TCA was added to the sample to make a final concentration of 20% and then left on ice for 30 minutes. The sample was then centrifuged at 8 000 g for 20 minutes at 4ºC and the pellet was washed twice with ice cold 100% acetone. Finally, the pellet was dried for 10 minutes and then re-suspended in equal volumes of water and protein loading buffer.

2.11 Translocation Kinetics

2.11.1 Pulse Chase protein labelling and immunoprecipitation.

Pulse chase labelling and immunoprecipitation was performed as described by Kwan *et al*., (2008) [152]. Samples were taken from the chase at 0, 10, 20 and 30 minutes and after immunoprecipitation they were analysed on a SDS page gel and visualised using Fluorography

2.12 *In vitro* translocation

2.12.1 Isolating Halophilic inverted membrane vesicles

Inverted membrane vesicles (IMV) are sealed membranes with an orientation opposite to that of intact cells and can be used for *in vitro* translocation of radiolabelled preAmyH. These IMVs were isolated as described by Kwan *et al*., (2008) [152].

2.12.2 *In vitro* translation and translocation

In vitro translation allows for the synthesis of radiolabelled preAmyH while *in vitro* translocation provides the means for the protein to translocate into the IMVs. Both procedures were performed as described by Kwan *et al*., (2008) [152].

2.13 Protein Assays

2.13.1 Amylase assays

Amylase activity was usually measured using the Phadebas amylase assay from Magle Life Sciences. One tablet was re-suspended in 10 ml of Buffer (5 mM $CaCl_2$, 4 M NaCl and 50 mM Tris pH 7.8) and then 1 ml of this suspension was added to 10 µl of purified AmyH or to the 210 µl of refolded AmyH solution. Unless stated, the reaction was left for one hour at 37ºC and then stopped by the addition of 1 ml of 0.5 M NaOH. Samples were then centrifuged at 14 000 g for 10 mins and absorbance was measured at OD_{620}.

The Phadebas amylase test could only be used for measuring the activity of purified AmyH. The more sensitive starch-iodine assay was used for measuring the activity of secreted AmyH and followed the protocol by Hutcheon *et al*., (2005) [150]. The only difference was that 100 µl of enzyme solution was added to the reaction and was then incubated for 2 hours.

2.13.2 Azocasein assay

Protease activity was measured using the azocasein assay. 150 µl of the enzyme solution was added to 250 µl of Buffer (4 M NaCl, 5 mM$CaCl_2$, 50 mM Tris-HCL pH 6.5 and 0.2% azocasein). The reaction was incubated overnight at 45°C and then terminated with 1 ml of 10% TCA. Samples were then put on ice for 30 minutes to allow undigested proteins to precipitate. Aggregated protein was removed by centrifugation for 10 minutes at 13 000 g. The supernatant was then mixed in a 1:1 ratio with 0.5 M NaOH and absorbance was measured at OD_{440}.

2.14 Bioinformatics

2.14.1 Determining Tat substrates

Open reading frames identified in the ongoing *H. volcanii* sequencing project were obtained from Dr Jonathan Eisen (UC Davis Genome Center). *H. volcanii* Tat substrates were identified using TatFind (http://signalfind.org/tatfind.html). The resulting proteins were analysed with TMHMM (http://www.cbs.dtu.dk/services/TMHMM) to filter out proteins with multiple membrane spanning domains.

To identify the *S. solfataricus* and *S. tokodaii* Tat substrates, the Pedant database (http://pedant.gsf.de/) was used in conjunction with SignalP (http://www.cbs.dtu.dk/services/SignalP/). The pattern search was x (1,35)-R-R-x-[FYVLIW]-[VLIMF]. X represents any amino acid, (1,35) indicates that this motif occurs in the first 35 amino acids of the sequence and the [ ] shows the different amino acids that can occur at this position. SignalP uses a Hidden Markov model and a Neural network. For sulfolobus sequences the models for Gram positive bacteria were used.

2.14.2 Sequence Logos

Sequence logos were produced using the SeqLogo software (http://imed.med.ucm.es/Tools/seqlogo.html).

2.14.3 Calculating hydrophobicity

The hydrophobicity of the H-domain was calculated by assigning each amino acid a value equivalent to that in the Kyte-Doolittle scale (http://www.expasy.ch/tools/pscale/Hphob.Doolittle.html). Both the total and average hydrophobicity of H-domains were calculated using these values.

Chapter 3: The *Sulfolobus* Twin Arginine Translocation pathway

3.1 The organisation of the *Sulfolobus* Tat complex

3.1.1 Introduction

A bioinformatic analysis of the *S. solfataricus, S. tokodaii* and *S. acidocaldarius* genomes revealed that they all contain a *tat* operon consisting of three genes. The two smaller genes encode the proteins $TatA_1$ and $TatA_2$ although it is not obvious if they are TatA or TatB proteins due to the similarities in structure. TatA and TatB in bacteria can be distinguished by TatA having a phenylalanine prior to the conserved glycine while TatB has a proline after this glycine [23]. As mentioned in the introduction, analysis of the *S. solfataricus* and *S. tokodaii* tat proteins hinted at the possibility that $TatA_1$ and $TatA_3$ are TatA proteins since they have this conserved phenylalanine. Meanwhile, there is less similarity to $TatA_2$ which lacks both the conserved phenylalanine and the proline. This of particular interest since so far there has been no TatB homologue discovered in archaea. Finally, the third gene encoding the larger protein has been found to be a homologue of the TatC protein. *S. solfataricus* also has a second smaller *tat* operon encoding $TatA_3$ and a TatC protein. Bioinformatics does not reveal enough information on these operons so experimental data are needed to understand the role of these proteins in Tat dependent transport.

Bolhuis *et al*., (2001) [54] purified the *E. coli* Tat complex *via* a Strep tag II attached to the C-terminal of TatC. This revealed that when in detergent the complex is around 600 kDa with TatB and TatC binding in a 1:1 ratio. It also showed that TatA was present within the complex but the majority was not purified because the protein formed separate homo-oligomeric complexes. Our goal was to purify the *S. solfataricus* and *S. tokodaii* Tat components in a similar manner to allow biochemical characterisation of the *Sulfolobus* Tat complexes.

3.1.2 Expression of the *S. tokodaii* and *S. solfataricus tat* genes

Expression of the *S. tokodaii* and *S. solfataricus tat* operons were first attempted in *E. coli*. Dr. Albert Bolhuis kindly provided the plasmid constructs $pSA_1A_2C_1$, pSA_3C_2 and $pTA_1A_2C_1$ shown in Figure 3.1. These constructs contain the *S. solfataricus* and *S. tokodaii* Tat operons downstream of the arabinose promoter of pBAD24. A Strep tag II was also incorporated onto the C-terminus of the TatC proteins to allow for detection and purification.

E. coli JM109 cells were transformed with $pSA_1A_2C_1$, pSA_3C_2 and $pTA_1A_2C_1$ and grown aerobically in LB supplemented with ampicillin and arabinose. When an OD_{600} of 1 was reached, cellular proteins were separated by SDS-PAGE and Western analysis was performed. Unfortunately the Strep tag II was not detectable which indicates that the TatC proteins were not or very poorly expressed (data not shown).

Sulfolobus genomes have a very low G-C content of approximately 35% [153]. It is possible that the *Sulfolobus tat* genes were not expressed because *E. coli* might not be able to transcribe this A-T rich DNA. To verify this, the presence of mRNA was tested. To this purpose, *E. coli* JM109 ($pSA_1A_2C_1$) and *E. coli* JM109 (pSA_3C_2) were grown to an OD_{600} of 0.8 where arabinose was added and allowed to grow for a further 2 hours. RT-PCR was then performed which revealed the presence of both the *S. solfataricus* $tatA_1A_2C_1$ and $tatA_3C_2$ mRNA (Figure 3.2).

The efficiency of translating the Tat operons would be greatly reduced if *E. coli* lacks the tRNAs used to transcribe the *Sulfolobus* A-T rich DNA. *E. coli* JM109 containing $pSA_1A_2C_1$, pSA_3C_2 and $pTA_1A_2C_1$ were transformed with a 2nd plasmid, pAYC-RIL. This plasmid encodes arginine, isoleucine and leucine codons that express tRNAs rare in *E. coli*. Transformed cells were grown to an OD_{600} of 0.8 where arabinose was added and allowed to grow for a further two hours. A Western blot was performed on the cells to detect the Strep tag II but the TatC protein could still not be identified (data not shown).

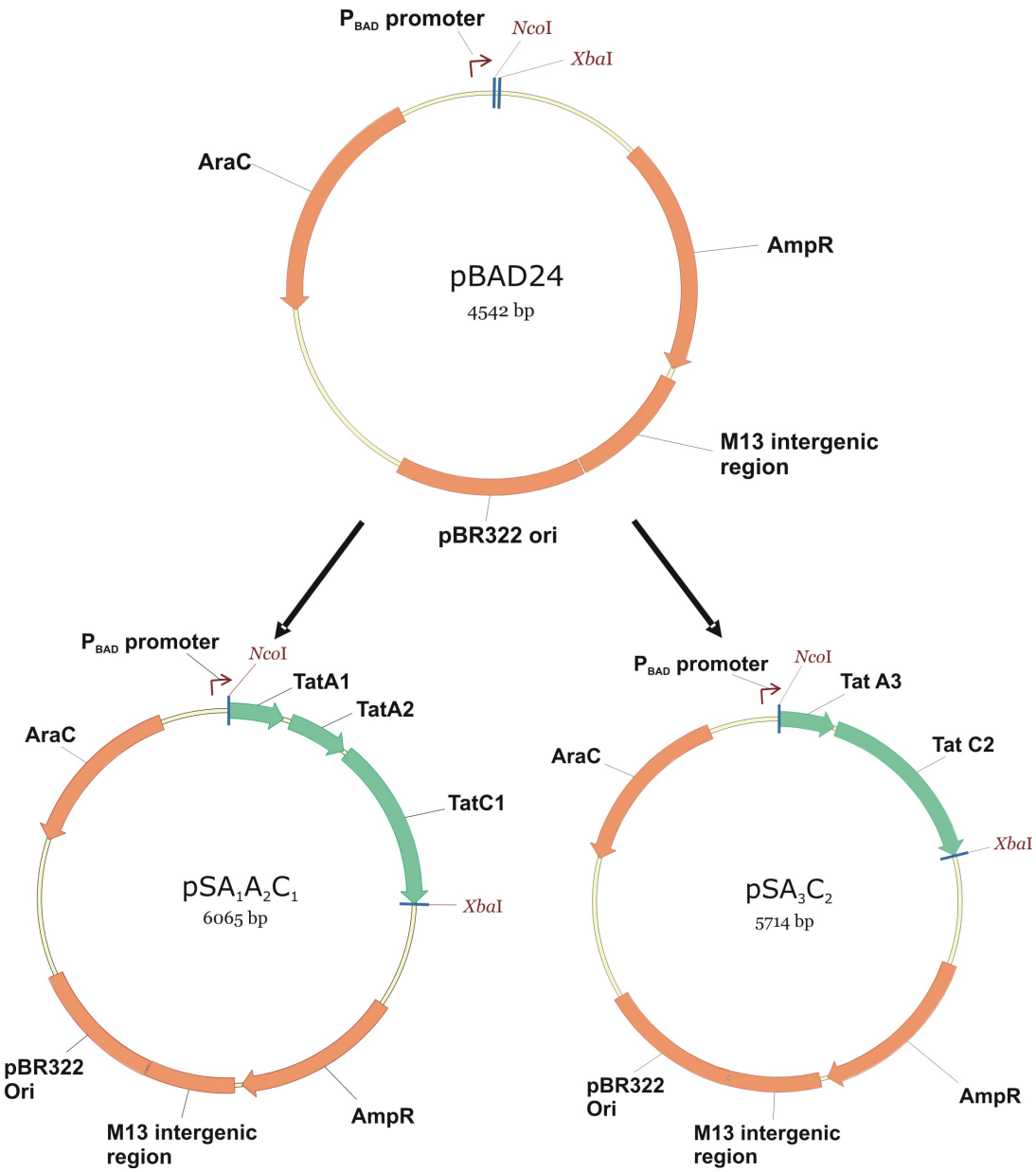

Figure 3.1 Construction of $pSA_1A_2C_1$ and pSA_3C_2. The *S. solfataricus* $A_1A_2C_1$ and A_3C_2 operons were cloned into the *Nco*I, *Xba*I site of the cloning vector pBAD24. A Strep tag was also incorporated onto the C-terminus of the $TatC_1$ and $TatC_2$ genes. The resulting vectors were denoted $pSA_1A_2C_1$ and pSA_3C_2. The *S. tokodaii* operon was cloned in the same way and named $pTA_1A_2C_1$. AmpR encodes a *β-lactamase* providing ampicillin resistance and AraC is a regulatory gene which encodes a cyclic AMP receptor protein required for L-arabinose utilization in *E. coli*.

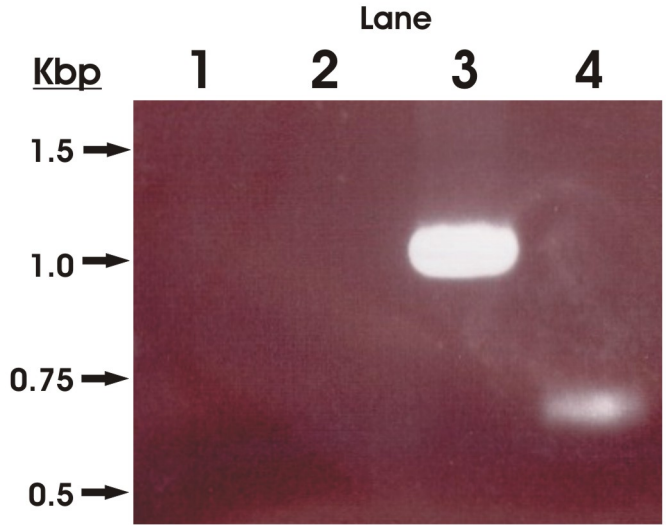

Figure 3.2 RT-PCR of the *S. solfataricus* $A_1A_2C_1$ and A_3C_2 mRNA in *E. coli* JM109. RT-PCR was performed on *E. coli* ($pSA_1A_2C_1$) and *E. coli* (pSA_3C_2) and the samples were run on an agarose gel. Lane 1 and 2 show the negative controls of *E. coli* ($pSA_1A_2C_1$) and *E. coli* (pSA_3C_2) respectively. These controls missed out the reverse transcription stage so that PCR was performed directly on the extracted mRNA, this confirms that there were no chromosomal contaminants. Lane 3 and 4 show the results of RT-PCR performed on *E. coli* transformed with $pSA_1A_2C_1$ and pSA_3C_2 respectively.

The *S. solfataricus* TatC proteins could not be detected in the experiments described above, and for that reason it was investigated whether expression of the *S. solfataricus tat* genes affected the growth of *E. coli* cells. To that purpose, *E. coli* JM109 ($pSA_1A_2C_1$) and *E. coli* JM109 (pSA_3C_2) were grown for 6 hours with and without the presence of arabinose (added after 3 hours growth). The absorbance was measured every hour and results are shown in Figure 3.3.

Results revealed that expression of either operon immediately inhibited cell growth. A reason for this could be that the *S. solfataricus* Tat proteins interact with the *E. coli* cell membrane and compromise membrane integrity by, for instance, forming complexes that are leaky for ions or protons. This could be because of interaction with the membrane by itself, or through interaction with *E. coli* Tat proteins. If the latter would be the case, the use of an *E. coli* mutant lacking its native Tat components would help in expression of the *Sulfolobus tat* genes. For that reason the *E. coli* mutant DADE (*ΔtatABCDE*) [46] was transformed with $pSA_1A_2C_1$and pSA_3C_2 and grown to an OD_{600} of 0.8 before being induced with arabinose for two hours.

Unfortunately cell growth was still inhibited immediately after induction, and the Strep tag II could not be detected on a Western Blot (data not shown). This indicates that is was not the interaction with the *E. coli* Tat components that was causing the problem.

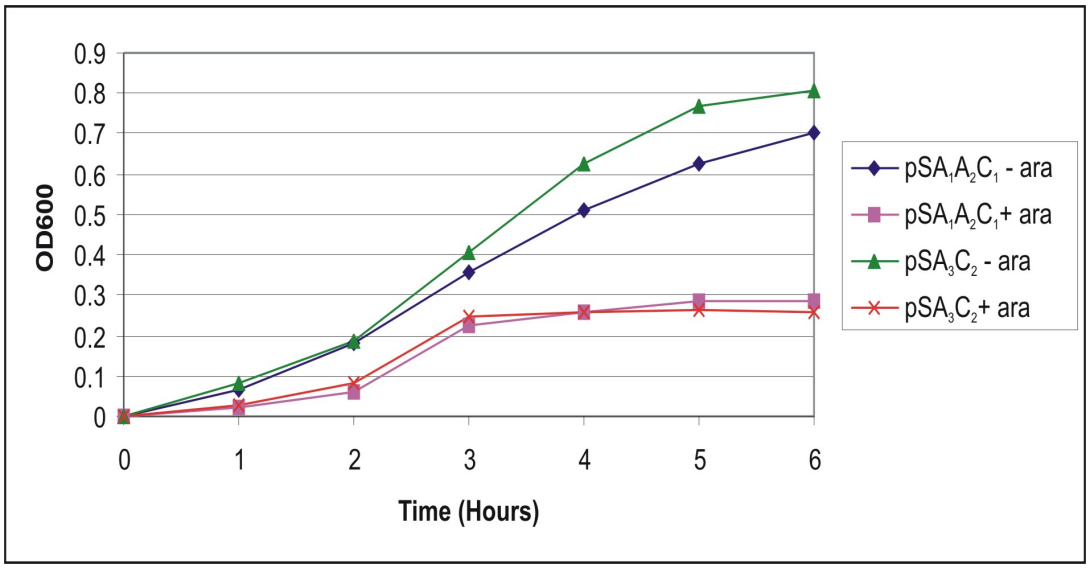

Figure 3.3 A 6 hour aerobic Growth Curve of *E. coli* JM109 transformed with $pSA_1A_2C_1$ and pSA_3C_2 in the presence/absence of arabinose. Cultures were grown in LB supplemented with ampicillin and after 3 hours arabinose was added. Absorbance (OD_{600}) readings were measured every hour for 6 hours. Data for the graph is shown in Appendix A.

3.1.3 Expression of the individual Tat complex proteins of *S. solfataricus*

The next step was to try and express the *Sulfolobus* Tat proteins individually in *E. coli*. The separate *S. solfataricus tat* genes were amplified using the primers in Table 3.1. These primers also incorporated the restriction sites *Nco*I (5') and *Hin*dIII (3'), as well as a sequence encoding a hexahistidine (His) Tag fused to the proteins C-termini.

Primer	Forward Primer	Reverse Primer
S.SoTatA1	AAA**CCATGG**CTTTAACAAACCC TTAGATTGGATCATAATACTAG TAGTAATTG	AAAAAAGCTTTTAAAA**AAGCTT**T TAgtgatggtgatggtgatgTTG ATTCTGCTTAGATTTCTTTAA
S.SoTatA2	AAAA**CCATGG**CAATTAGCAACT TAAGTGACTTTTTAG	AAAA**AAGCTT**CTAgtgatggtga tggtgatgGTTCTTTCCGTCACT AGCTTT
S.SoTatC1	AAAA**CCATGG**CAATATATTTCA TGCTGGGTTTAGAAT	AAAA**AAGCTT**CTAgtgatggtga tggtgatgTTTTTCGAACTGCGG GTGGCTCCAA
S.SoTatA3	AAAA**CCATGG**CAGCTATAGATA ATCCTACTGATATAATTATAAT CTTAGTGGTTATTG	AAAA**AAGCTT**TCAgtgatggtga tggtgatgGGCCTTCTTGCTTTG CTTTAGTGC
S.SoTatC2	AAAA**CCATGG**CAACAGGTTATG AATGGATTAATATAGATCCATT ACAAA	AAAA**AAGCTT**CTAgtgatggtga tggtgatgTTTTTCGAACTGCGG GTGGCTCCA

Table 3.1 The primers used to amplify the individual *S. solfataricus* Tat proteins for expression in *E. coli*. The restriction enzyme sites are underlined and are in bold (*Nco*I/*Hin*dIII). The His Tag is noted as small red letters in the Reverse primer. All primers are annotated 5'-3'.

The amplified genes were then ligated behind the arabinose promoter of the cloning vector pBAD24 using the *Nco*I/*Hin*dIII restriction sites. The resulting vectors were denoted pSA1, pSA2, pSC1, pSA3 and pSC2 and encode the *S. solfataricus* $TatA_1$, $TatA_2$, $TatC_1$, $TatA_3$ and $TatC_2$ proteins respectively. The vector pSA1 is shown in Figure 3.4, the other four vectors were constructed in a similar manner.

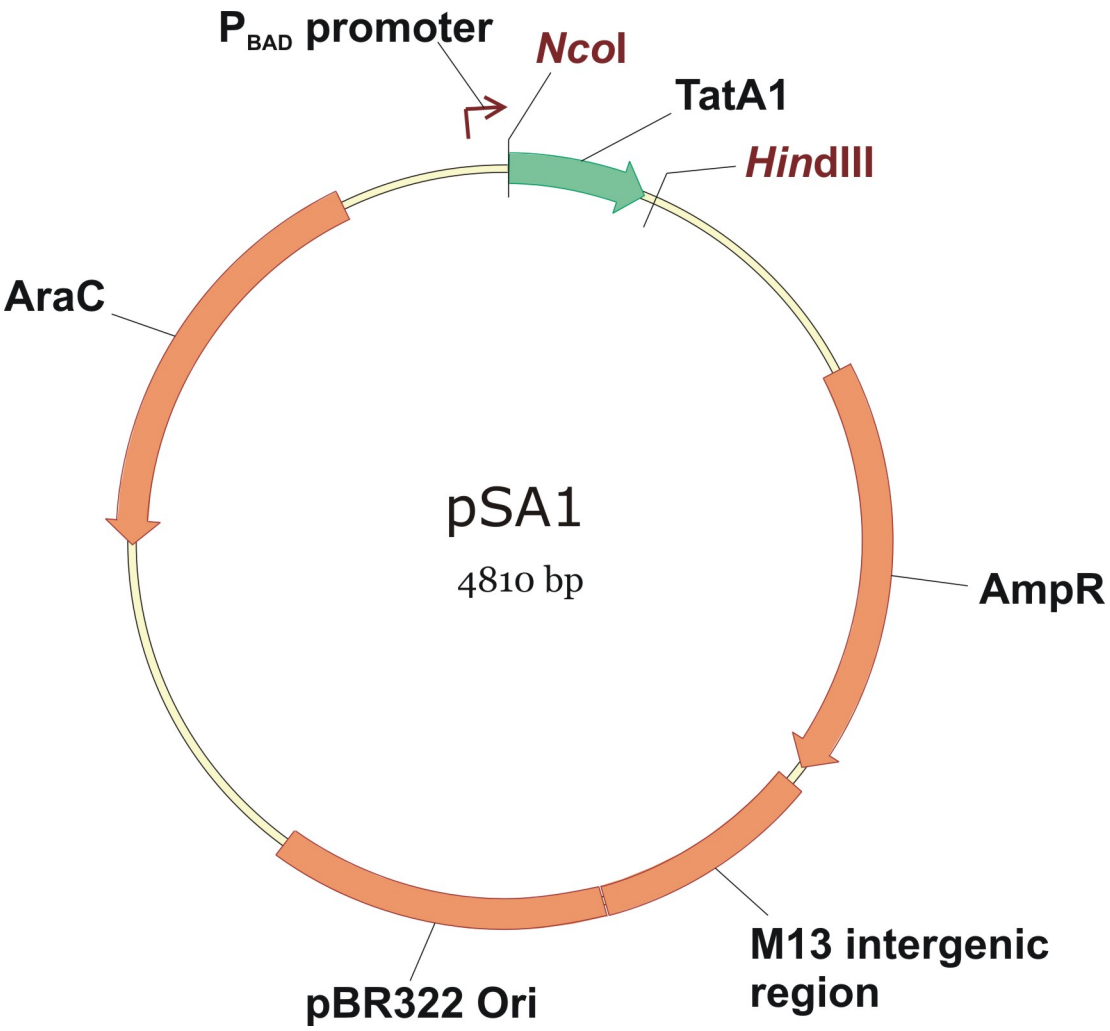

Figure 3.4 Vector pSA1. The *S. solfataricus TatA*$_1$ gene was ligated into the *NcoI, Hin*dIII site of pBAD24. A His tag was also incorporated onto the C-terminus of the protein. The resulting construct formed pSA1. AmpR encodes a *β-lactamase* providing ampicillin resistance and AraC is a regulatory gene which encodes a cyclic AMP receptor protein required for L-arabinose utilization in *E. coli*.

E. coli JM109 cells were transformed with these *S. solfataricus* expression vectors and grown until an OD_{600} of 0.8. Arabinose was then added and cultures were grown for a further 2 hours. Cells were harvested and a Western blot was performed on the cellular and membrane fractions to detect the proteins C-terminal His tag.

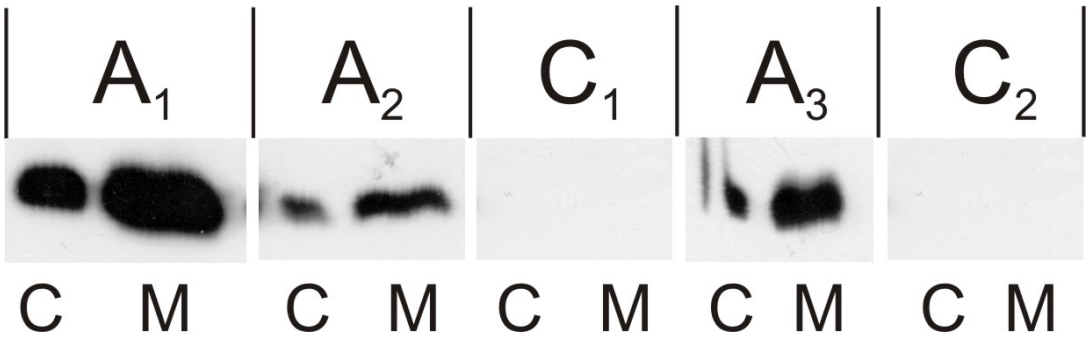

Figure 3.5 Expression of the individual *S. solfataricus* Tat protein. Lanes from left to right show the *E. coli* cytoplasmic (C) and membrane (M) fractions when cells were transformed with pSA1, pSA2, pSC1, pSA3 and pSC2 respectively. All cultures were grown in 5 ml LB broth with ampicillin and arabinose for 8 hours and after fractionation a Western blot was performed. Double the amount of the sample containing $TatA_2$ was loaded onto the SDS Page gel.

Figure 3.5 clearly shows the presence of all three *S. solfataricus* TatA proteins in both the cytoplasmic and membrane fractions of *E. coli*. For unknown reasons the $TatA_2$ was expressed at significantly lower levels than the other two Tat proteins. Unfortunately, similar to previous experiments the expression of the TatC proteins inhibited growth (Figure 3.6) and the His tag could not be detected on a Western blot (data not shown).

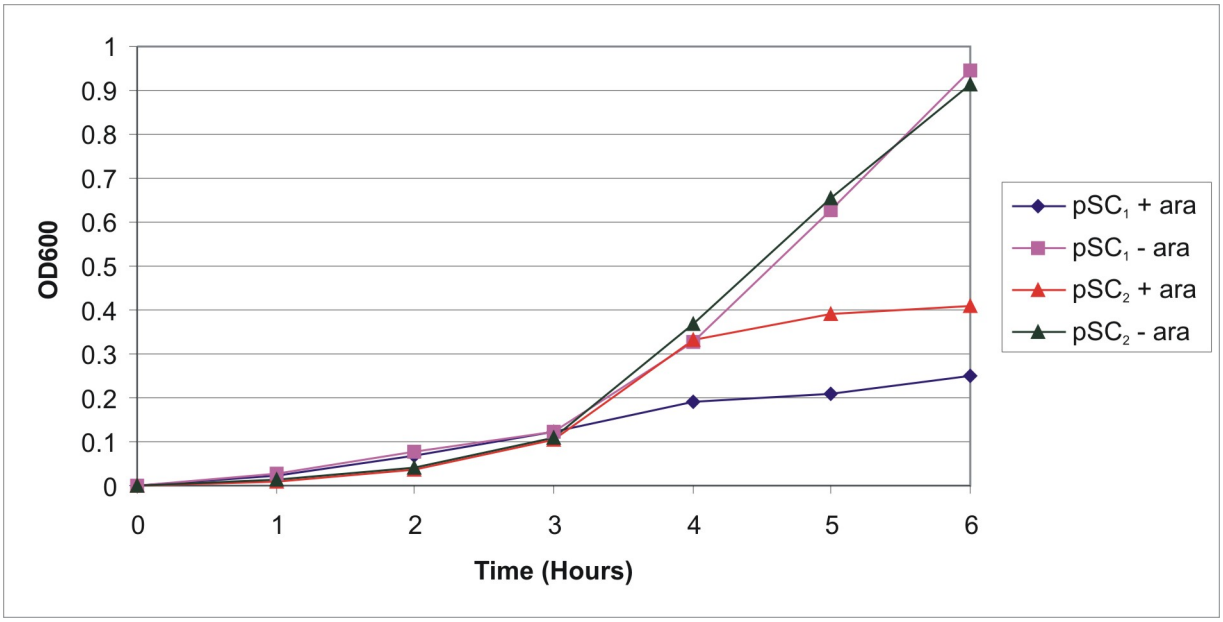

Figure 3.6 A 6 hour aerobic Growth Curve of *E. coli* JM109 transformed with pSC_1 and pSC_2 in the presence/absence of arabinose. Cultures were grown in LB supplemented with ampicillin and after 3 hours arabinose was added. Absorbance (OD_{600}) readings were measured every hour for 6 hours. Data is shown in Appendix A.

3.1.4 Purification and structural studies of $TatA_3$

Section 3.1.3 showed that the *Sulfolobus* TatA proteins can be expressed in *E. coli* therefore allowing them to be purified for further studies. *E. coli* JM109 (pSA3) was grown in 8 litres of LB broth supplemented with ampicillin and arabinose. When an OD_{600} of 1 was reached the cells were lysed and passed through a HisPur Cobalt column. The $TatA_3$ protein bound to the column was then purified and the separate elution fractions were run on the Coomassie stain shown in Figure 3.7.

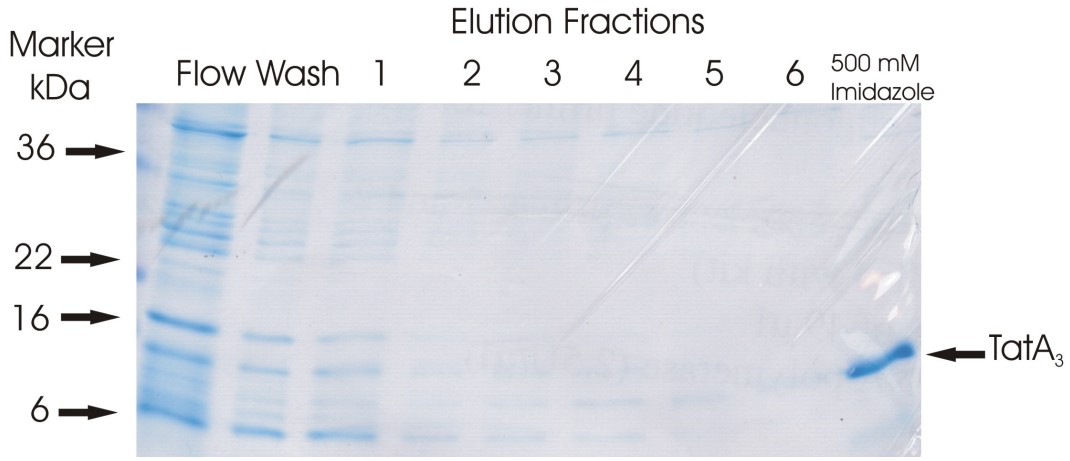

Figure 3.7 Purification of $TatA_3$. $TatA_3$ was purified using a HisPur Cobalt column and the resulting fractions were run on a Coomassie stain. Lane 1: The flow through containing the initial unbound cellular proteins. Lane 2: The wash fraction containing the loosely bound proteins washed from the column. Lane 3-9: 1ml elution fractions containing 100 mM Imidazole. Lane 10: Final Elution fraction containing 500 mM Imidazole.

The elution fractions containing 150 mM imidazole (as recommended by the manufacturer) were not capable of displacing the $TatA_3$ from the column. This proved advantageous because the elution buffer removes all background proteins. A very pure sample of $TatA_3$ was then eluted by using buffer containing 500 mM imidazole as shown in Lane 9.

Purified $TatA_3$ was then sent to Covalab UK to raise anti-$TatA_3$ polyclonal antibodies for further research. Unfortunately the antibody produced high background on a Western blots making it impossible to distinguish the $TatA_3$ protein (data not shown). Purification of the antibody would increase its specificity but due

to time constraints this experiment has not been done. Crystallisation of $TatA_3$ was also attempted in collaboration with Professor S. Iwata and J. Andrell (Imperial College). Unfortunately, this was unsuccessful since $TatA_3$ precipitated under all conditions that were tested.

Finally, gel filtration was performed on purified $TatA_3$ to determine if it forms monomeric or homo-oligomeric structures. A Western blot of the resulting fractions revealed that $TatA_3$ forms complexes ranging between 60 and 135 kDa (Figure 3.8).

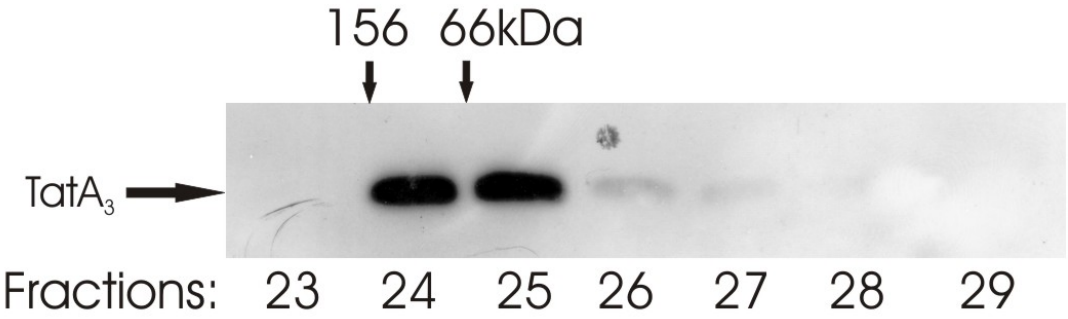

Figure 3.8 Gel filtration of Tat A_3. Gel filtration was performed on $TatA_3$ and then a Western blot was performed on the resulting fractions using Invitrogen's anti-His antibodies (mouse). $TatA_3$ is shown to be in fractions 24 and 25. Also noted are the fractions that aldolase (156 kDa) and BSA (66 kDa) elute from so that they can be used for size comparison.

3.1.5 The compatibility of the *S. solfataricus* $TatA_1$ and $TatA_3$ proteins in the *E. coli* Tat system

The TMAO assay is a non-quantitative method to show the presence of a functioning Tat pathway in *E. coli* [53, 68]. It utilises methyl viologen which is blue in colour when reduced but looses its colour if oxidised. To perform the assay the cytoplasmic, membrane and periplasmic fractions of an *E. coli* culture are run on a native protein gel. The gel is stained in buffer containing methyl viologen and then transferred to a buffer containing TMAO. If present, the *E. coli* Tat substrate TorA will oxidise the methyl viologen causing a clearing on the gel.

TorA is a periplasmic protein, although with native PAGE significant amounts of activity are also found in the cytoplasm and membrane. Importantly, TorA is not translocated to the periplasm in the absence of a fully functional Tat pathway. The TMAO assay can therefore be used to determine if the *S. solfataricus* $TatA_1$ and $TatA_3$ proteins are compatible with the *E. coli* Tat system. pSA1 and pSA3 were transformed into a *E. coli* Δ*tatA/E* mutant. Cultures were then grown aerobically to an OD_{600} of 1 in LB broth containing ampicillin and arabinose. The cytoplasmic, membrane and periplasmic fractions were then fractionated and analysed on a native PAGE gel. The TMAO assay was then performed and results are shown in Figure 3.9.

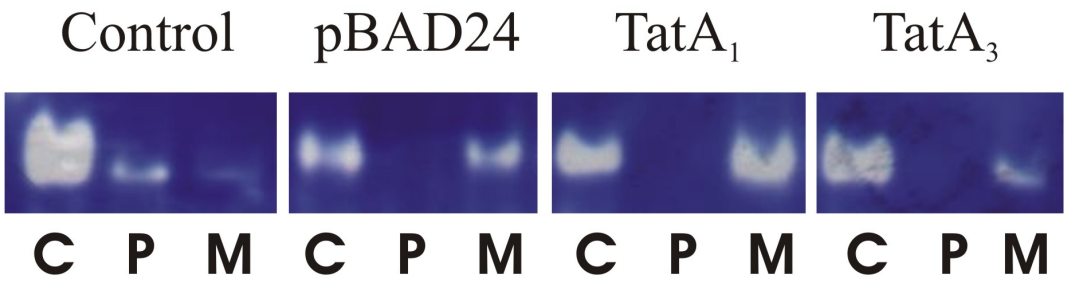

Figure 3.9 TMAO assay of *E. coli* Δ*tatA/E* transformed with pSA1 and pSA3. The first 3 lanes are the positive control and show the cytoplasmic (C), periplasmic (P) and membrane (M) fraction of *E. coli* MC4100. The next 3 lanes are the fractionated samples from the negative control, *E. coli* MC4100 Δ*tatA/E* (pBAD24). Lanes 7-12 show the separate fractions of an *E. coli* MC4100 Δ*tatA/E* mutant transformed with pSA1 ($TatA_1$) and pSA3 ($TatA_3$).

The positive control of the TMAO assay represents a functional Tat pathway in *E. coli* and as expected TorA is present in all three fractions. The negative control is the *E. coli* Δ*tatA/E* strain which lacks a functional Tat system therefore TorA is only present in the cytoplasm and the membrane. For *E. coli* Δ*tatA/E* transformed with pSA1 and pSA3 there is no clearing in the periplasmic fraction which indicates that TorA has not translocated and that $TatA_1$ and $TatA_3$ are not capable of complementing the deletion of *tatA* and *tatE*.

3.1.6 Expression of the *S. solfataricus* $TatA_1A_2C_2$ and $TatA_3C_2$ proteins in the native organism

The overproduction of proteins in *Sulfolobus* spp. is a relatively new technique that is still being developed by a number of different groups. The method by Albers *et al*., (2006) [131] utilises the viral vector pMJO5 as an expression vector. Due to the size of pMJO5 (20 Kb) the cloning is done in two stages. First, the desired gene is ligated into the entry vector pMZ1 behind the arabinose promoter. This section is then digested out and ligated into pMJO5. The completed construct can then be transformed into the *S. solfataricus* M16 strain which lacks *pyrEF*, a uracil synthesis gene. This deletion is complemented by the *pyrE* and the *pyrF* gene on pMJO5 allowing successful transformants to grow in the absence of uracil. After transformation, expression of the gene of interest is achieved by the addition of 0.4% D-arabinose.

The *S. solfataricus* $tatA_1A_2C_1$ and $tatA_3C_2$ operons were amplified using the primers in Table 3.2. The primers incorporated the restriction sites *Nco*I and *Bam*HI, therefore allowing the amplified gene to be ligated into pMZ1. The new constructs were then digested with *Avr*II and *Eag*I and the smaller fragment was ligated into viral vector to form pEXPA1A2C1 and pEXPA3C2. The construction of the pEXPA1A2C1 expression plasmid is shown in Figure 3.10.

	Forward Primer	Reverse Primer
S.so A1A2C1 to insert into pMZ1	AAA**CCATGG**CTTTAACAAACCC TTACGATTGGATCATAATACTA GTAGTAATTG	AAAA**GGATCC**AATTCTTTCCAA AGTTTTCTGGGATTCTTTC
S.so A3C2 to insert into pMZ1	AAA**CCATGG**CTATAGATAATCC TACTGATATAATTATAATCTTA GTG	AAAA**GGATCC**AGATTCCCTTTT CCTCTCTACTAA

Table 3.2 The primers used to amplify the *S. solfataricus* operons for ligation into pMZ1. The restriction sites *Nco*I and *Bam*HI are underlined and in bold. All primers are annotated 5'-3'.

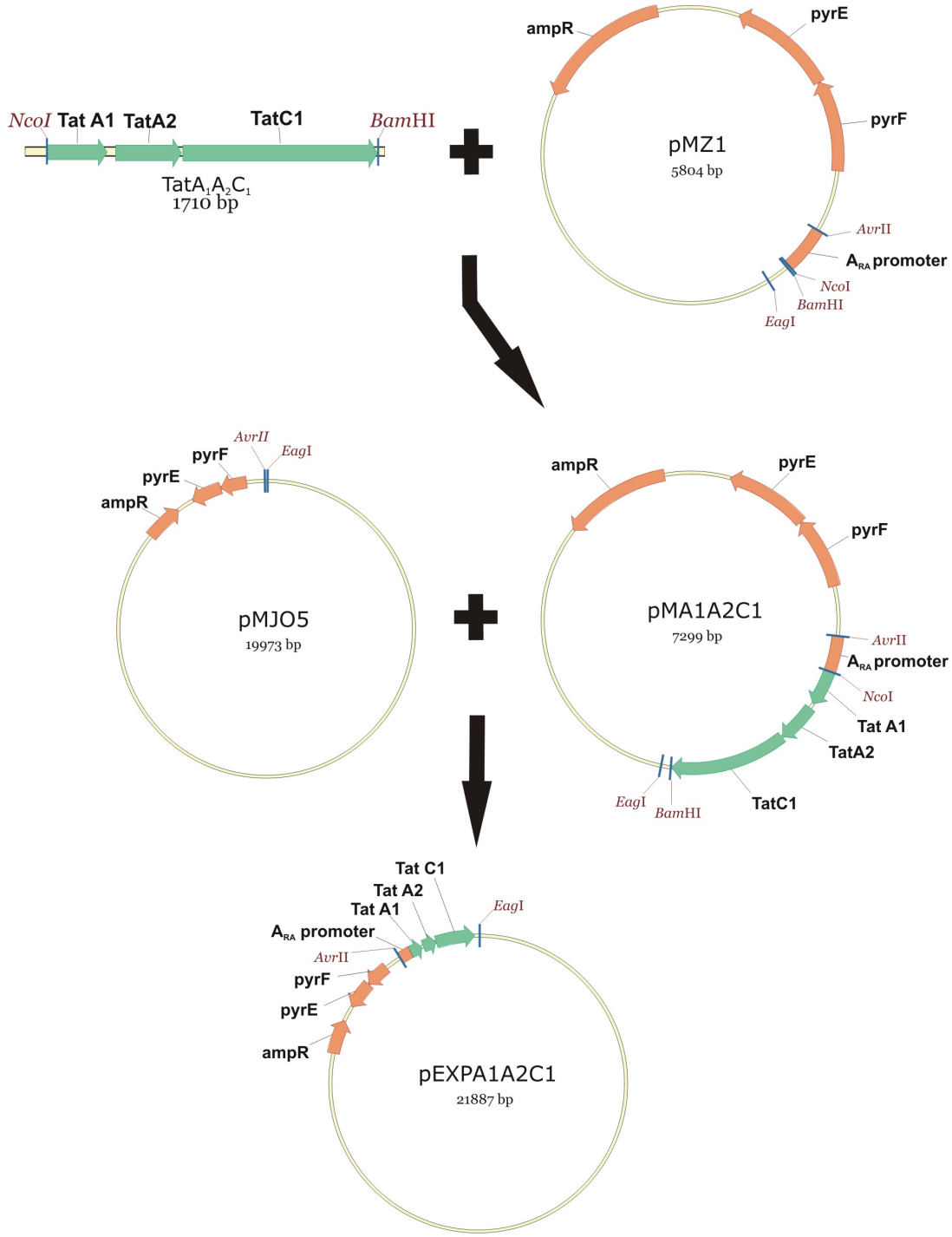

Figure 3.10 The construction of the Sulfolobus expression vector pEXPA1A2C1. The initial *S. solfataricus tatA$_1$A$_2$C$_1$* operon is ligated into *Nco*I/*Bam*HI restriction sites of the entry vector pMZ1 to form pMA1A2C1.The smaller fragment of pMA1A2C1 after an *Eag*I/*Avr*II double restriction digest is then ligated into the viral vector pMJO5. The resulting plasmid forms the final expression vector pEXPA1A2C1. pEXPA3C2 containing the *S. solfataricus tatA$_3$C$_2$* operon was made in a similar manner. AmpR encodes a β-lactamase providing ampicillin resistance and *pyrEF* is a uracil synthesis gene.

S. solfataricus M16 (pEXPA1A2C1) and *S. solfataricus* M16 (pEXPA3C2) were incubated in 0.1% tryptone Brock medium containing 10 µg/ml of uracil. Cells were allowed to regenerate for two days and then transferred into Brock media lacking uracil. Unfortunately cultures did not reach an OD_{600} greater then 0.3 indicating that transformation was unsuccessful. Repeated attempts yielded similar results even when *S. solfataricus* M16 was transformed with empty vector. Since *S. solfataricus* transformations were successful in Section 3.2.2 using the knockout vector it indicates that there was a problem with the viral vector itself and not the transformation protocol.

3.2 The role of the *S. solfataricus* $tatA_1A_2C_1$ and $tatA_3C_2$ operons

3.2.1 Introduction

S. solfataricus differs from *S. tokodaii* and *S. acidocaldarius* because it contains a second tat operon consisting of a *tatA* and *tatC* gene. No other archaeon has been found to contain two *tat* operons but it is a common feature among many Gram positive bacteria. Pop *et al*., (2002) [70] have shown that in *B. subtilis* the first Tat complex can only translocate YwnB while the second translocates PhoD. Not only is the second complex specific but it is only transcribed when PhoD is required under phosphate starvation. Section 3.2 aims to determine if *S. solfataricus* also uses its two operons to provide specificity to the Tat substrates that they translocate. However, it is also possible that the second Tat operon is just a redundant copy of the first and does not actually have a function within the organism.

A new technique has been developed by Albers and Driessen (2007) [132] that allows *S. solfataricus* genes to be deleted from the genome. It works by amplifying 700 bp of DNA either side of the gene of interest. These regions are ligated into the non replicating vector pET2268 so that they flank the *lacS* gene (Figure 3.11A). This gene encodes a β-glycosidase that is required for growth on media containing lactose as the sole carbon source [147]. The resulting construct is transformed into *S. solfataricus* PBL2025 (Δ*lacS* strain) and cells can only grow on lactose minimal Brock medium if pET2268 has incorporated into the chromosome *via* homologous recombination (Figure 3.11B). A second recombination event is then required for the gene of interest to be lost from the chromosome.

This recombination can either occur at the same point as the previous crossover, therefore restoring the wildtype strain, or it can occur at the opposite flanking region, allowing the gene of interest to be lost and resulting in a successful knockout (Figure 3.11). Unfortunately, there is no selectable marker to select for the knockout recombination so cultures are continually re-inoculated until the gene of interest is lost naturally. Section 3.2.2 aims to use this technique to produce *S. solfataricus*

mutants that lack the tat operons. It is hoped that this will reveal if the Tat pathway is essential and also provide an insight into the function of the two separate Tat complexes.

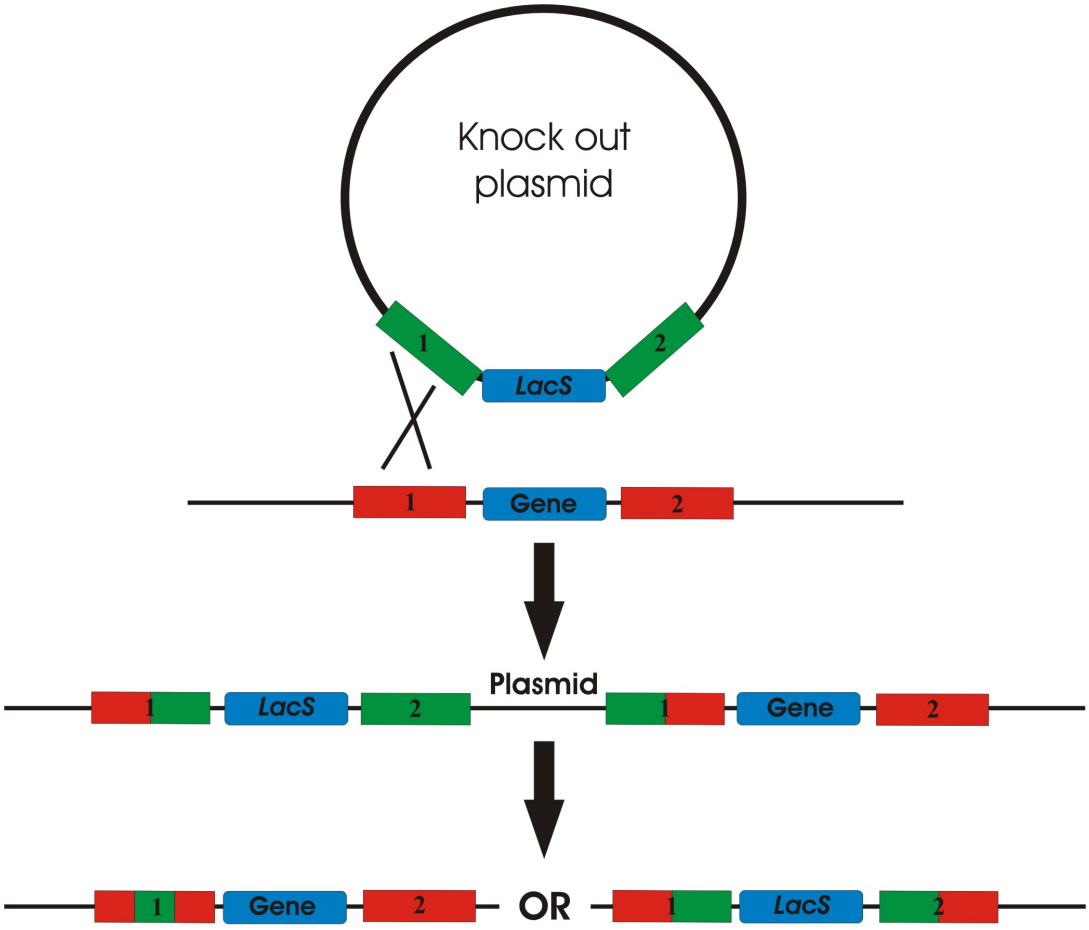

Figure 3.11 A scheme for creating a successful gene knockout in *S. solfataricus*. First the plasmid integrates into the chromosome via a single recombination event that occurs between homologous regions of the two pieces of DNA. There are then two possibilities for where the second recombination event can occur. Either within the same region as the first therefore maintaining the wildtype strain or within the second homologous region creating a successful knockout strain. The 1 and 2 are the 700bp up and down stream regions of the gene of interest that have been ligated into the knockout vector. The 1 and 2 are the corresponding homologous chromosomal up and down stream regions. The black cross indicates the region where the first crossover has occurred.

As mentioned previously, the *B. subtilis* second *tat* operon is only transcribed under phosphate starvation [70, 73]. It is therefore a possibility that one of the two *S. solfataricus tat* operons is also only transcribed under specific conditions. Section

3.2.3 looks at the mRNA expression of the Tat operons and their substrates under different growth conditions. It is hoped that by doing so it will reveal if there is some form of regulation occurring.

3.2.2 Producing *S. solfataricus* Δ*tatA$_1$A$_2$C$_1$* and Δ*tatA$_3$C$_2$* knockout mutants

The primers in Table 3.3 were used to amplify the 700 bp flanking regions of the *S. solfataricus tatA$_1$A$_2$C$_1$* and *tatA$_3$C$_2$* operons. These DNA fragments were then ligated into the *Kpn*I/*Nco*I and *Bam*HI/*Not*I restriction sites of pET2268. The resulting constructs formed pKOA1A2C1 and pKOA3C2 which are shown in Figure 3.12 and are designed to knockout the *S. solfataricus tatA$_1$A$_2$C$_1$* and *tatA$_3$C$_2$* operons respectively.

	Forward Primer	Reverse Primer
S.so $A_1A_2C_1$ Upstream flank	ATAT**GGTACC**ATTCATATTT AGGACTTACTGTGTTAAAAC TTT	TATA**CCATGG**TGTTTTCATTCG AGTGGGAGGT
S.so $A_1A_2C_1$ Downstream flank	ATAT**GGATCC**AGAATTTAGA AAAATTACGTGATTCGGTTT ATGAAGTTTT	TATATATA**GCGGCCGC**ATTATG CTATTTATAAAAACGTTCCCGC CAGCG
S.so A_3C_2 Upstream flank	ATATATAT**GCGGCCGC**CCTC ATGTACATCCTCTATTGTTA GTCTTTTACCCATA	TATA**GGATCC**ATCTATCATAAT TTACCCTCCCGCCCCA
S.so A_3C_2 Downstream flank	ATAT**CCATGG**GAATCTTAAG TTATAAATAGTAGTTATCTT TATAATCAACACT	TATA**GGTACC**TTCCGATATTAT AAAAAATAGTAAAAATTGTGAT ACACCCACTAA

Table 3.3 The primers used to amplify the 700bp Upstream and Downstream flanking regions of the *S. solfataricus tatA$_1$A$_2$C$_1$* and *tatA$_3$C$_2$* operons. The restriction sites *Kpn*I,*Bam*HI, *Nco*I and *Not*I are shown in bold and are also underlined. All primers are annotated 5'-3'.

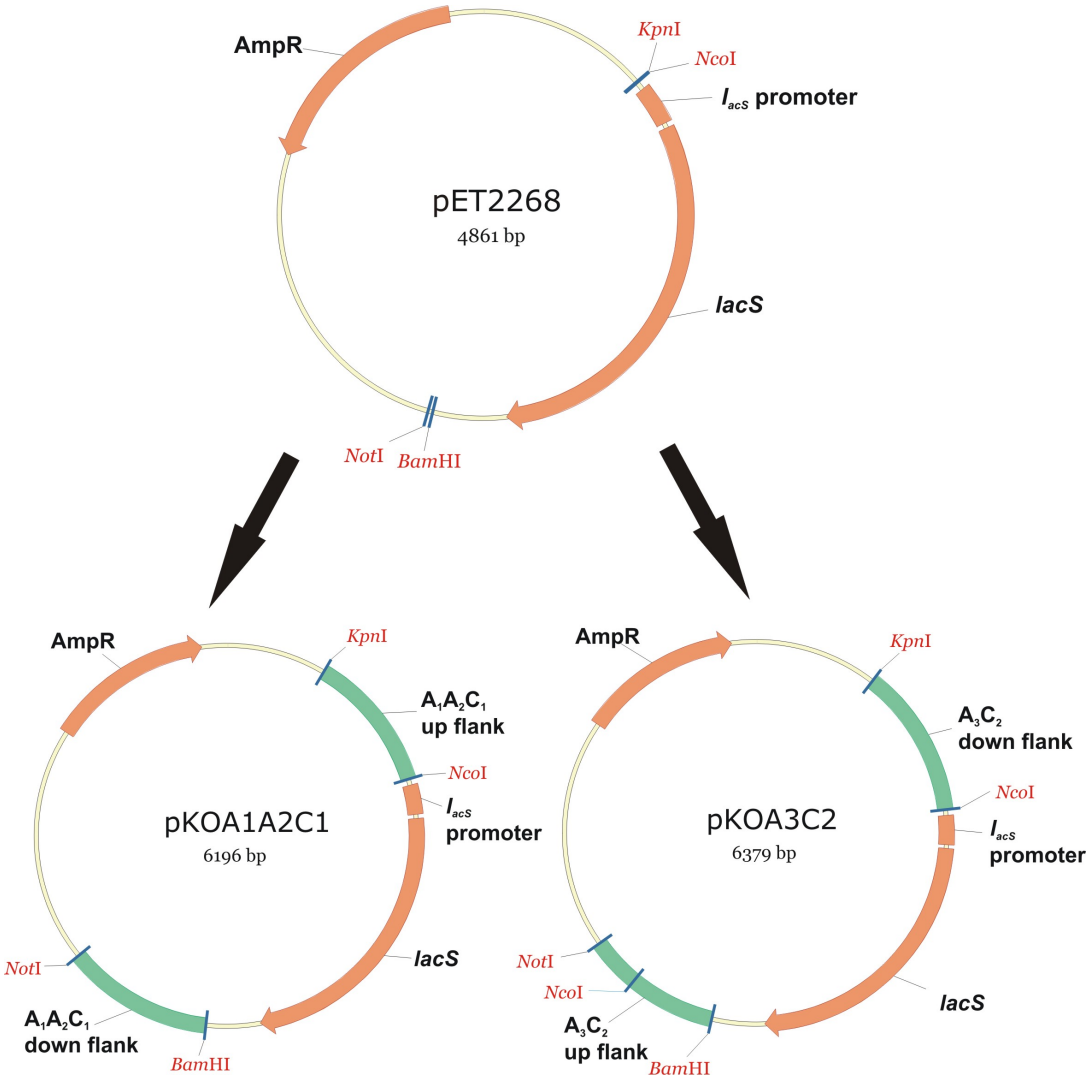

Figure 3.12 The construction of the *Sulfolobus* knockout vectors pKOA1A2C1 and pKOA3C2. A schematic of the *E. coli*/*S. solfataricus* shuttle vector pET2268 with the 700 bp flanking regions of the *S. solfataricus* tat$A_1A_2C_1$ and tatA_3C_2 operons ligated into the *Kpn*I/*Nco*I and *Not*I/*Bam*HI restriction sites. These new constructs form pKOA1A2C1 and pKOA3C2 respectively. AmpR encodes a *β-lactamase* providing ampicillin resistance and *lacS* encodes a β-glycosidase that is required for growth on media containing lactose as the sole carbon source

S. solfataricus PBL2025 cells were transformed with pKOA1A2C1 and pKOA3C2 and successful recombination was indicated by the resulting blue colonies when sprayed with X-gal. These colonies were re-inoculated into lactose Brock medium and allowed to grow to an OD_{600} of 0.8. Chromosomal DNA was extracted and PCR was performed to determine if the tat operons were still present. The primers that were used are shown in Table 3.4.

	Forward Primer	Reverse Primer
S.so$A_1A_2C_1$ check	AGTTAGAGGAGTTAAAGAAA TCTAAGC	GTACGCCAATTCTCTCAAATGT TCTAT
S.soA_3C_2 check	GGAAGATCAATGGGGGAGTT CAA	CTGCAATACTATCGTAAAATGA CGGAT
lacS check	TCTACTTTTAAGATCCTTGA ATATTTCCCTG	**For $A_1A_2C_1$ transformation –** ACTATCTAACCTCCCACTCGAA **For A_3C_2 transformation -** GTGTTGATTATAAAGATAACTA CTATTTATAACTTAAG

Table 3.4 The primers used to determine if the *S. solfataricus* knockout mutants had been successful. The S.so$A_1A_2C_1$ check and S.soA_3C_2 check primers were designed to amplify 439 bps of the S. *solfataricus* $A_1A_2C_1$ and A_3C_2 genes. The lacS check primers were designed to amplify a 657 bp region of the S. *solfataricus* $A_1A_2C_1$ mutant DNA or 639bps of the S. *solfataricus* A_3C_2 mutant DNA. The region that is amplified contains a section of the *lacS* gene and the flanking region of the Tat operon therefore can be used to confirm plasmid integration. All primers are annotated 5'-3'.

Unfortunately, after 5 rounds of re-culturing, PCR still revealed the presence of both the *lacS* gene and the $tatA_3C_2$ operon (Figure 3.13). Similar results were also found for *S. solfataricus* transformed with the $tatA_1A_2C_1$ knockout vector (data not shown).

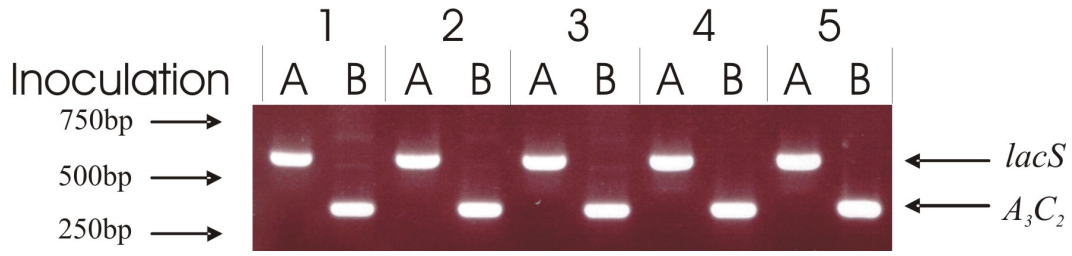

Figure 3.13 PCR to determine if the *S. solfataricus* $tatA_3C_2$ operon or the *lacS* gene is present. The DNA gel shows the PCR performed on chromosomal DNA of *S. solfataricus* transformed with the $tatA_3C_2$ knockout plasmids. The 5 rounds of re-inoculation each contain two PCR reactions, the first is for a 439 bp section of the $tatA_3C2$ operon (A) and the second is for the 639 bp *lacS* gene (B).

It is possible that double recombination was not occurring because there is no selectable marker to force the plasmid to pop out of the chromosome. Several

problematic *E. coli*, *H. volcanii* and *Methanosarcina acetivorans* mutants were achieved by transforming linear DNA rather then the complete knockout plasmid [154-156]. This is because linear DNA can only integrate into the chromosome *via* a double cross over event. To date, no publications have shown that this technique has been used with *Sulfolobus*, but if successful then it would allow the *lacS* gene to be used as the selectable marker for double recombination. The plasmid pKOA1A2C1 and pKOA3C2 were digested with *Kpn*I and transformed into *S. solfataricus* pBL2025. Unfortunately, after 14 days no growth was detected on lactose medium for both transformations.

3.2.3 Transcription of the *S. solfataricus* Tat operons and their Tat substrates

Transcription of the Tat operons was analyzed when *S. solfataricus* was grown on either rich or chemically defined medium. To this purpose cells were grown to an OD_{600} of 0.5 in Brock media containing tryptone or glucose, respectively. RT-PCR was then performed to detect the $tatA_1A_2C_1$ and $tatA_3C_2$ mRNA. The resulting DNA was analysed on an agarose gel and results are shown in Figure 3.14. This revealed that the $tatA_1A_2C_1$ and $tatA_3C_2$ mRNA is present when grown under both carbon sources. The negative control confirms that there was no chromosomal contamination, which would have caused a false positive signal.

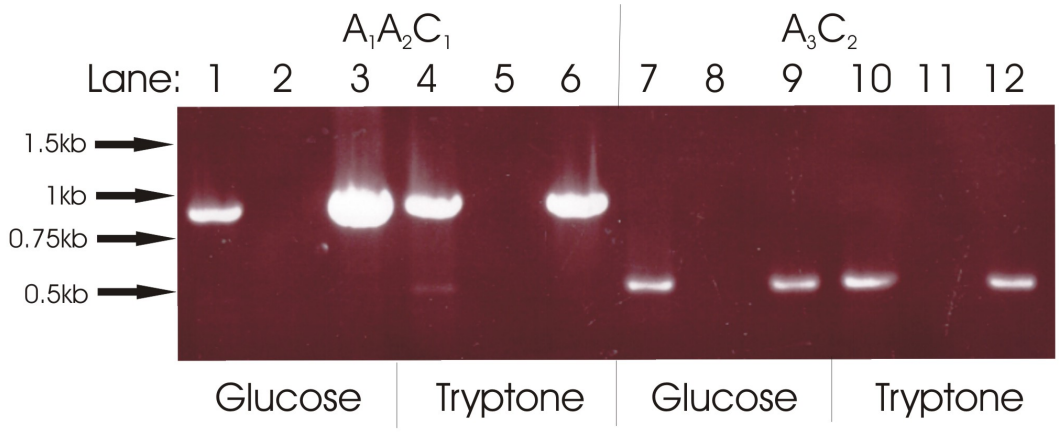

Figure 3.14 RT-PCR of the *S. solfataricus* $tatA_1A_2C_1$ and $tatA_3C_2$ operons. Lanes 1, 4, 7 and 10 show RT-PCR of the $tatA_1A_2C_1$ and $tatA_3C_2$ mRNA from *S. solfataricus* grown on glucose or tryptone media. Lanes 2, 5, 8 and 11 show the negative controls where no RT-PCR was performed prior to PCR. Lanes 3, 6, 9 and 12 show the positive controls where a PCR reaction to detect the Tat operons was performed on *S. solfataricus* chromosomal DNA.

Finally, the transcription of the predicted Tat substrates was analysed in rich media conditions. *S. solfataricus* was grown in tryptone Brock media until an OD_{600} of 0.5. RT-PCR was then performed to detect the mRNA of six predicted Tat substrates. Results are shown in Figure 3.15 which revealed the presence of mRNA for all six genes.

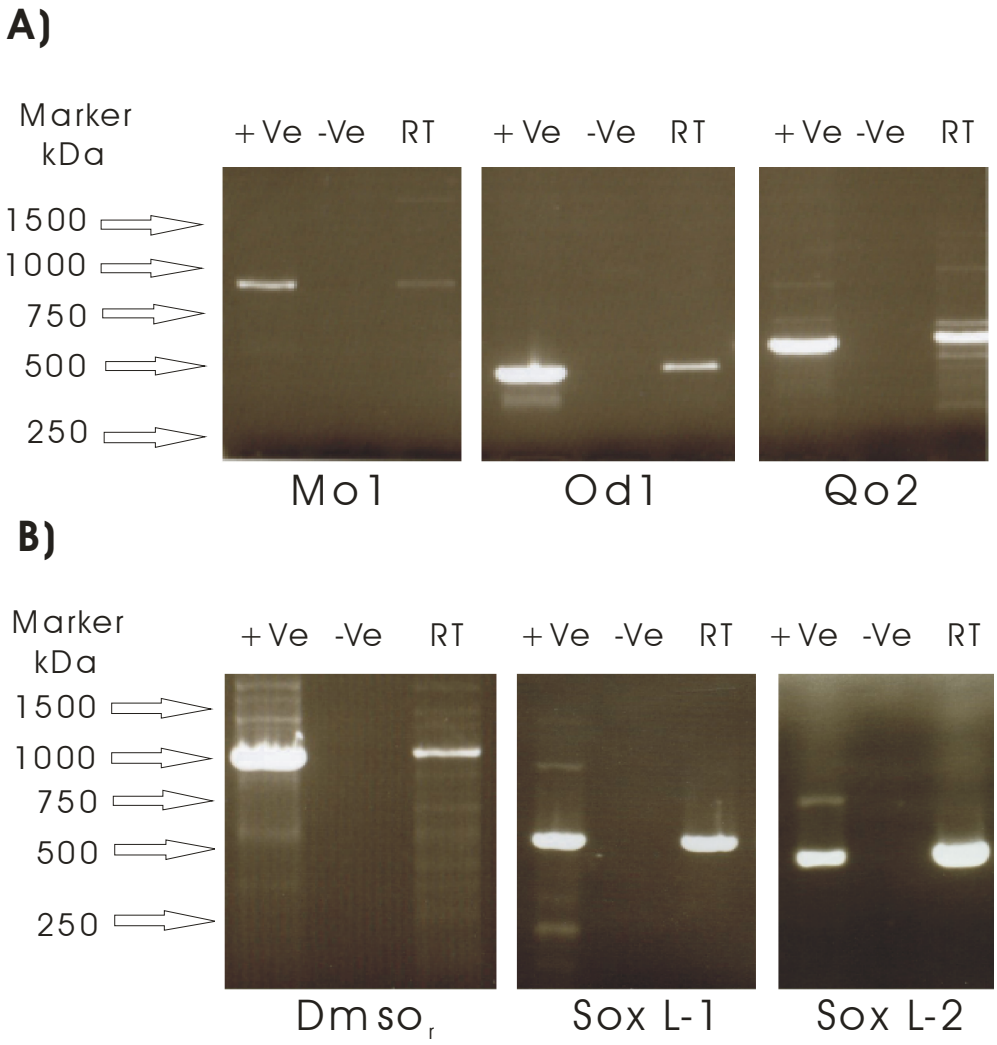

Figure 3.15 RT-PCR for the predicted Tat substrates in *S. solfataricus*. *S. solfataricus* was grown in rich tryptone media and then RT-PCR was performed for the Tat substrates. The shorthand for the *S. solfataricus* Tat substrates are: Mo1 = Molybdopterin oxidoreductase, $DMSO_r$ = DMSO reductase, Od1 = hypothetical oxidoreductase, SoxL-1 = Rieske iron sulphur protein 1. SoxL-2 = Rieske iron sulphur protein 2 and Qo2 = Quinol oxidase-2,, the Rt lanes show the RT-PCR for the tat substrate mRNA when grown on 0.4% tryptone media. The negative control shows the PCR on the mRNA when no RT-PCR has been performed. The positive control is the same PCRs but performed on chromosomal DNA.

3.3 Discussion

Bolhuis *et al*., (2001) [54] previously purified the *E. coli* TatC protein by incorporating a Strep-tag II onto its C-terminus. By doing so it also allowed the co-purification of the TatA and TatB proteins. This provided the basis of our knowledge into the interactions between TatA, TatB and TatC in *E. coli*. Chapter 3.1 looked at performing a similar experiment by purifying the Tat complexes of *S. tokodaii* and *S. solfataricus*. The Strep-tag II attached onto the C-terminal of TatC was not detectable when the *S. tokodaii* and *S. solfataricus* operons were expressed in *E. coli*. It is likely that this is because *Sulfolobus* spp. are so vastly different from *E. coli* on a molecular and biochemical level.

Firstly it was possible that *E. coli* does not have sufficient amounts of the tRNAs required for efficient expression of the *Sulfolobus* A-T rich DNA [99, 153]. RT-PCR revealed that the operons mRNA is present so transcription is occurring. However, this technique is not quantitative so it is possible that there not enough mRNA for efficient protein expression. Another issue is that *Sulfolobus* utilises several codons that occur at a low frequency in *E. coli* [99]. It has been shown that this can considerably reduce the translation of *Sulfolobus* Tat genes especially if these codons cluster within the ORF [157]. When *E. coli* was complemented with a plasmid encoding several of these rare tRNAs the transcription of the operons was still not detectable [158].

A growth curve revealed that the expression of the Tat components immediately inhibited *E. coli* growth. It was hoped that expression in the *E. coli* mutant DADE would overcome any problems if the *Sulfolobus* Tat complex would interact with the *E. coli* components. However growth was still inhibited by induction and the TatC Strep-tag II was not detected, indicating that the protein is likely interacting with the membrane itself in a toxic manner. Similar results were found by Prisco *et al*., (1995) [159] who tried to express a *S. solfataricus* putative membrane protein in *E. coli*. They suggested that because the *Sulfolobus* membrane proteins have evolved to interact with a very specific lipid composition, it has meant that they are not compatible with the *E. coli* membrane [160].

The Tat components were expressed individually in the hope that it would have less impact on cell integrity. This was true for the TatA proteins which could now be detected. The production of the *S. solfataricus* TatC proteins was still not possible and a growth curve revealed that their expression was still toxic to the cell. Unfortunately this will make expression of the *Sulfolobus* TatC proteins difficult to achieve in *E. coli.* It may be possible that another expression system such as using Gram positive bacteria or yeast may be more successful. For example, the thermostable β-galactosidase from *S. solfataricus* was successfully expressed in *Saccharomyces cerevisiae* [161].

$TatA_3$ was purified so that further structural and biochemical studies could be performed. Gel filtration revealed that $TatA_3$ does forms homo-oligomeric complexes that range between 60 and 135 kDa. It must be remembered that the detergent will also contributes to this mass, and the number of TatA proteins in the 60-135 kDa complexes is therefore not clear [162]. This result concludes that $TatA_3$ does form homo-oligomeric complexes, but it seems likely that their size will be different when expressed in the native organism. The six predicted *S. solfataricus* Tat substrates range from 11.5 to 157.9 kDa thus it can be speculated that the Tat complex can accommodate proteins between these size. However these proteins could translocate as dimers or as subunits of a larger complex, which would in turn lead to the requirement of a larger pore. However, to my knowledge there is no literature available to suggest that this is the case.

So far, crystallisation of the *E. coli* TatA protein has been problematic, probably due to the variety of different sized complexes that are purified (C. Robinson, personal communication). Since $TatA_3$ only ranges from 60-135 kDa when expressed in *E. coli* it was hoped that it would prove easier to form individual crystals. Samples of pure $TatA_3$ were sent to Professor S. Iwata and J. Andrell at Imperial collage so that they could attempt crystallisation. Unfortunately they were unsuccessful since $TatA_3$ precipitated under all conditions investigated.

Recent research has found that specific heterogeneous TatA, TatB and TatC proteins can function in the *E. coli* Tat system. For example, *Helicobacter pylori* TatA and *Pseudomonas syringae* TatA and TatB can complement an *E. coli* mutant lacking the

corresponding components [53, 163]. The Tat pathway of a *E. coli* Δ*tatABCDE* strain can even be restored by the expression of the *Streptomyces coelicolor* or the *Magnetospirillum Magneticum tat* operon [164]. Research by Xiong *et al*., (2007) [164] found that expression of the *S. solfataricus* $tatA_1A_2C_1$ and $tatA_3C_2$ operons would not complement a *E. coli* Δ*tatABCDE* strain. Unfortunately, they did not confirm expression of the operons which is important since my results suggests that TatC is toxic. For that reason, $TatA_1$ and $TatA_3$ were expressed in *E. coli* and the TMAO assay was performed. This confirmed that these two *S. solfataricus* TatA proteins were not capable of complementing an *E. coli* Δ*tatA/E* mutant. This was not surprising since there is such a large contrast in the cytoplasmic environment and the membrane composition found within the two organisms.

The expression of the *S. solfataricus* Tat operons was also attempted using the native organism. It was hoped that this would overcome the problems found when working with *E. coli*. This experiment would also provide a much more accurate view on the protein interactions and reveal if there are novel proteins that are involved. Unfortunately no growth was seen for *S. solfataricus* M16 transformed with the expression plasmid pEXPA1A2C1 or pEXPA1A2C1 in Brock media lacking uracil. This was also the case for *S. solfataricus* M16 transformed with the empty vector, pMJO5 which indicates a problem with the expression vector or with the transformation procedure. Unfortunately, due to time constraints no more work was performed on the expression of these *Sulfolobus* Tat proteins.

The Tat pathway is essential for the translocation of many important co-factor binding proteins such as oxidoreductases and Rieske iron sulphur proteins. This means that if the pathway is blocked then it is likely to affect the organism's phenotype. For example, if the *E. coli* Tat pathway is blocked then the organism is no longer capable of anaerobic respiration on specific substrates and also become impaired of cell separation during cell division leading to a chain formation [26, 27]. This is because several of the Tat substrates are important components involved in the anaerobic respiratory chain, outer membrane assembly, and cell separation. By knocking out the *S. solfataricus* $tatA_1A_2C_1$ and $tatA_3C_2$ operons it was hoped that it would reveal if there are any changes in the organisms phenotype. This would give

an indication into the role of the two Tat operons and their importance in *S. solfataricus*.

Results revealed that only single recombination occurred when transforming *S. solfataricus* with pKOA1A2C1 and pKOA3C2. It is possible that the two operons are maintained in the chromosome because there is no selectable marker to force the double recombination event. This seems unlikely since Albers and Driessen (2007) [132] have successfully created several mutants using this technique. The other possibility is that the Tat pathway is actually an essential process in *S. solfataricus* and therefore is lethal to the organism if removed. The *S. solfataricus* Tat pathway is predicted to translocate three Rieske iron sulphur and three oxidoreductases. The Rieske iron sulphur proteins form essential subunits of the cytochrome bc-complex important in energy metabolism [112], and the oxidoreductase proteins will be involved in the electron transfer between two separate molecules. Therefore, it is conceivable that preventing these proteins from translocating could be lethal to the cell. The only other archaeal species where this system has been found to be essential is the haloarchaeon *H. volcanii* and possibly also *Halobacterium salinarum.* However, this was not surprising since the Tat system plays such an important role in Haloarchaea as it is involved in the secretion of the majority of proteins [119]. There has been some evidence that the Tat pathway is essential in the bacteria *Mycobacterium tuberculosis* as it is not possible to produce Δ*tatAC* knockout mutants. This organism encodes eleven predicted Tat substrates, four of them are phospholipase C proteins involved in liberating phospholipids as a carbon and energy source, several are possibly involved in scavenging for metal ions and the others have unknown functions [165].

A difference in the Tat pathway of *E. coli* and *B. subtilis* is that the *E. coli tat* genes are expressed constitutively while one of the *B.subtilis tat* operons is only expressed under phosphate starvation [26, 73]. *S. solfataricus* is similar to *B. subtilis* in that it is also has two separate Tat operons. It was therefore hypothesised that each operon would encode for Tat complexes with some form of specificity to particular Tat substrates. Although RT-PCR is not quantitative, it revealed that the two Tat operons and all of the predicted Tat substrates are expressed under all conditions tested. Finally, the substrate DMSO reductase is known in *E. coli* to act as a terminal reductase that reduces DMSO allowing growth by anaerobic respiration [166, 167].

Although, there is no known literature on *S. solfataricus* being able to grow anaerobically they do express DMSO reductase as a Tat substrate. Unfortunately, all attempts to grow the organism anaerobically using DMSO as the electron acceptor failed (data not shown). Similarly, Trimethylamine N-oxide (TMAO) which is a Tat substrate in *E. coli* can be reduced by TMAO reductase (TorA) as part of an anaerobic respiratory chain [168, 169]. Although blast searches indicate that *S. solfataricus* lacks a protein homologous to TorA, growth on TMAO under anaerobic conditions should be tested.

To conclude, only purification of the *S. solfataricus* TatA proteins was possible in *E. coli*, as the TatC proteins were toxic for *E. coli* cells. It was found that TatA forms homo-oligomeric complexes in *E. coli*, but TatA1 and TatA3 were not able to complement the *E. coli* Δ*tatA/E* mutant. Unfortunately, expression of the *tat* operons in *S. solfataricus* was not possible and it is suspected that this was due to problems with the viral vector. Finally, attempts to knockout the *S. solfataricus tat* operons failed, which could suggest that the Tat pathway is essential in *S. solfataricus.*

Chapter 4: The twin arginine motif of haloarchaeal Tat substrates

4.1 The use of bioinformatics to analyse the haloarchaeal Tat signal peptide

4.1.1 Introduction

The Tat signal peptide has been found to be the critical factor for folded proteins to translocate *via* the Tat pathway [44]. This is due to the highly conserved twin arginine motif [35], the hydrophobic region [32] and the positively charged Sec avoidance residues [34]. The twin arginine motif is found between the N-domain and the H-domain and has the consensus sequence (S/T)-R-R-x-F-L-K [22]. The twin arginines are nearly always present and the surrounding 5 amino acids occur at a high frequency.

Uniquely, due to the nature of their highly saline environment, haloarchaea utilise the Tat pathway as their major route of translocation [117, 118]. It is not known whether there are differences in the properties of the haloarchaeal Tat and Sec signal peptides, or whether there are differences between the signal peptides of Tat substrates from haloarchaea and non-halophilic organisms. Therefore a bioinformatics study into the frequency of the amino acids in the haloarchaeal Tat signal peptides may yield some interesting results.

It has also been shown that by replacing the H-domain of the *E. coli* Tat substrate TorA with a more hydrophobic region would cause the protein to be rerouted *via* the Sec pathway [32], which confirmed the importance in the hydrophobicity of the H-domain in directing bacterial proteins to the Tat or Sec system. It is therefore important to determine if there is also such a difference between the hydrophobicity of haloarchaeal Tat and Sec substrates.

4.1.2 The frequency of the amino acids from halophilic Tat signal motifs

The motifs of all the Tat substrates from *E. coli* and *Halobacterium* sp. NRC-1 were available from Bolhuis (2002) [170] and Robinson and Bolhuis (2001) [23]. The Tat substrates from *H. marismortui* and *N. pharaonis* were kindly provided by Hutcheon and Bolhuis (unpublished data) while the *H. volcanii* Tat substrates were determined using a combination of TatFind and TMHMM (the *H. volcanii* genome sequence was kindly provided by Jonathan Eisen, UC Davis Genome Centre (University of California)). The motifs were then aligned and the SeqLogo program was used to create sequence logos, the results are shown in Figure 4.1. The purpose of these logos is to provide a visual method for easier analysis into the frequency of amino acids within a group of protein sequences. For example, the position of the amino acid on the column shows its frequency compared to the other amino acids at this position, while the size of the amino acid shows proportionally how much more this amino acid occurs.

It appears that the Tat motifs of all five organisms analysed follows a similar frequency in their amino acids. Position 1 is mainly a serine or a threonine and position 4 is mainly a phenylalanine, valine or leucine. There is some difference at position 7 with *E. coli* having a strong frequency for lysine while the halophilic organisms have a preference towards alanine. Another clear difference is that, compared to the *E. coli* Tat substrates, haloarchaea have a much stronger preference for the leucine at position 6. Finally, it appears that a variety of amino acids can occupy position 5 with there being no obvious preference.

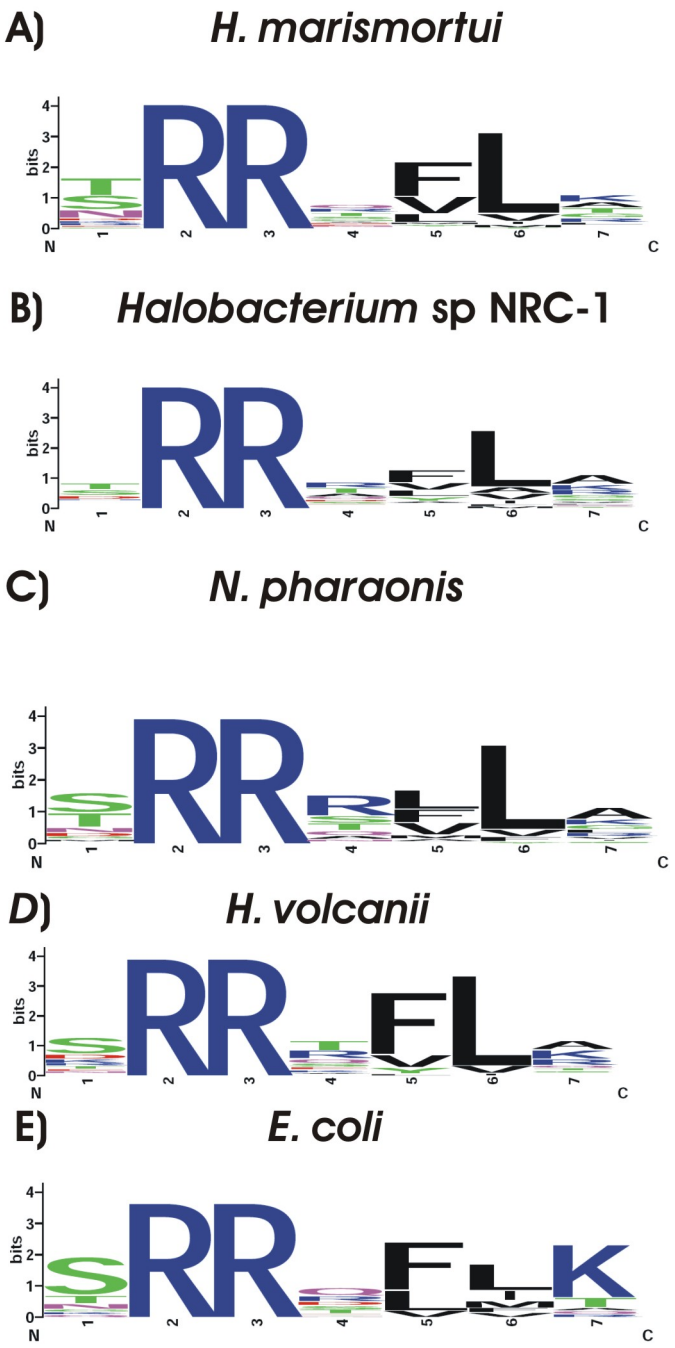

Figure 4.1 Sequence Logos of the Tat motifs from a variety of halophiles, alkaliphiles and *E. coli*. The Tat motifs of each organisms Tat substrates were aligned and then converted into SeqLogo so that the data could be analysed in the form of sequence logos. For the purpose of the Sequence Logos, the amino acid prior to the twin arginines has been designated as position 1 of the Tat Motif. A) The *H. marismortui* logo analysed the organisms 61 predicted Tat substrates B) The *Halobacterium* sp. NRC1 logo analysed the organisms 60 predicted Tat substrates C) The *N. Pharaonis* logo analysed the organisms 40 predicted Tat substrates D) The *H. volcanii* logo analyzed the organisms 40 predicted Tat substrates E) The *E. coli* logo analyzed the organisms 22 predicted Tat substrates.

4.1.3 The hydrophobicity of the H-domain

The sequences of the H-domain from the signal peptides of the Tat and Sec substrates of *E.coli* and *Halobacterium* sp. NRC-1 were available from Robinson and Bolhuis (2001) [23] and Bolhuis (2002) [170] respectively. Due to the quantity of *E. coli* Sec substrates, a random selection of 22 were chosen for analysis. The hydrophobicity of these H-domains was calculated using the Kyte-Doolittle scale and the hydrophobicity per amino acid was also determined to take into account the variation in H-domain length. The results are shown in Table 4.1 which revealed that although there is little difference in the length of the H-domain, it is more hydrophobic in the Sec substrates of *E. coli* and *Halobacterium* sp. NRC1 compared to their Tat counterparts.

	Length (sig. peptide)	**Length (H-domain)**	**Total hydrophob. (H-domain)**	**Average hydrophob. (H-domain)**
E. coli Tat substrates	35.91	16.91	23.60 $\pm$ 2.66	1.43 $\pm$ 0.21
E. coli Sec substrates	22.59	12.32	28.27 $\pm$ 4.18	2.32 $\pm$ 0.34
H. sp. NRC-1 Tat substrates	29.33	13.39	22.16 $\pm$ 7.55	1.70 $\pm$ 0.48
H. sp. NRC-1 Sec substrates	26.86	16.86	38.64 $\pm$ 8.88	2.35 $\pm$ 0.39

Table 4.1 The hydrophobicity of the H-domain from the Tat and Sec substrates of *Halobacterium* sp. NRC-1 and *E. coli*. The signal peptides form the *Halobacterium* sp. NRC-1 Tat and Sec substrates and the *E. coli* Tat and Sec (A selection of 22 proteins) substrates were analysed. The hydrophobicity of the H-domain was calculated using the Kyte-Doolittle scale (http://www.expasy.ch/tools/pscale/Hphob.Doolittle.html). The average hydrophobicity of the amino acids in the H-domain was calculated by dividing the hydrophobicity of the H-domain by the length of the H-domain. Data for the individual proteins is shown in Appendix B.

4.2 The use of the protease SptA to analyse the importance of its twin arginine motif for translocation via the Tat pathway

4.2.1 Introduction

The bioinformatic study described above suggests that the twin arginines play a critical role in the targeting and translocation of haloarchaeal Tat substrates as is seen with all other organisms. For example with *E. coli*, substituting the twin arginines of Tat substrates to twin lysines will block translocation and the same occurs if just the second arginine is mutated [36]. Mutation of the first arginine to a lysine does allow some translocation to occur although efficiency is greatly reduced [33]. It therefore important to determine if the twin arginines seen in the haloarchaeal Tat substrates are also an essential requirement for efficient translocation, especially considering such a large proportion of halophilic secretory proteins utilise this pathway.

Research into the twin arginine motif of halophilic Tat substrates has mainly been on the α-amylase from *Natronococcus amylolyticus* [171] and the serine protease (SptA) from *Natrinema* sp J7. SptA. SptA is the focus of this chapter and contains a strong Tat signal peptide as shown in Figure 4.2 [172].

Figure 4.2 The signal peptide of SptA. The region in red shows the Tat motif and the region in blue shows the hydrophobic region of the signal peptide. The arrow indicates the processing site.

Shi *et al*., (2006) [172] confirmed SptA as a Tat substrate by creating a twin lysine SptA mutant (R11K, R12K) and showing that translocation had ceased by using an SptA translocation assay. This assay works by expressing SptA in *H. volcanii* and allowing the protease to translocate into the media. If translocation is successful then the protease will degrade milk provided in the agar therefore forming a halo (Figure 4.3). Since this twin lysine mutation prevented translocation of SptA it meant that a halo was no longer produced.

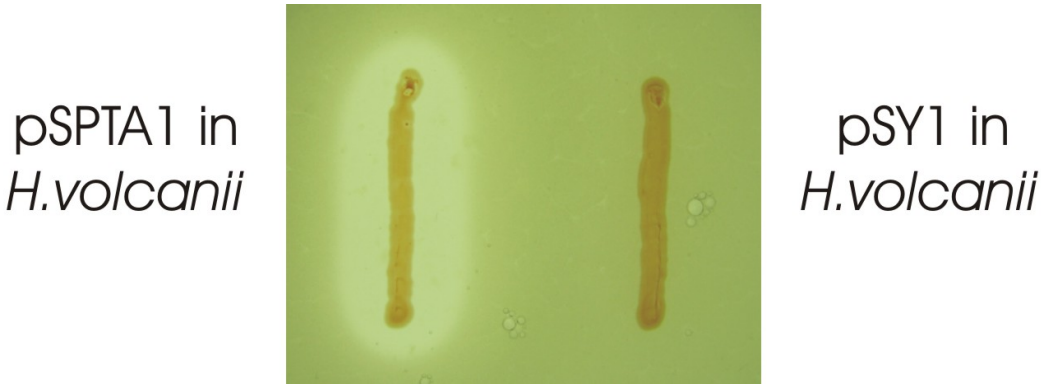

Figure 4.3 An SptA translocation assay. On the left is *H. volcanii* containing the SptA expression vector, pSPTA1. On the Right is *H. volcanii* containing the empty shuttle vector, pSY1. Colonies were allowed to grow for 3 days.

Shi *et al*., (2006) [172] have already shown that the twin arginines play an important role in the translocation of the halophilic Tat substrate, SptA. It would be interesting to use this SptA plate assay to investigate the importance of the individual arginines. It may be the case that like with *E. coli*, the presence of both arginines is not essential for translocation to occur.

4.2.2 The importance of the Twin Arginines for Tat translocation

Shi *et al*., [172] kindly provided the SptA expression plasmid pSPTA1 shown in Figure 4.4. It was constructed by ligating SptA behind the B_{op} promoter of the pSY1 *E. coli*/*H. volcanii* shuttle vector.

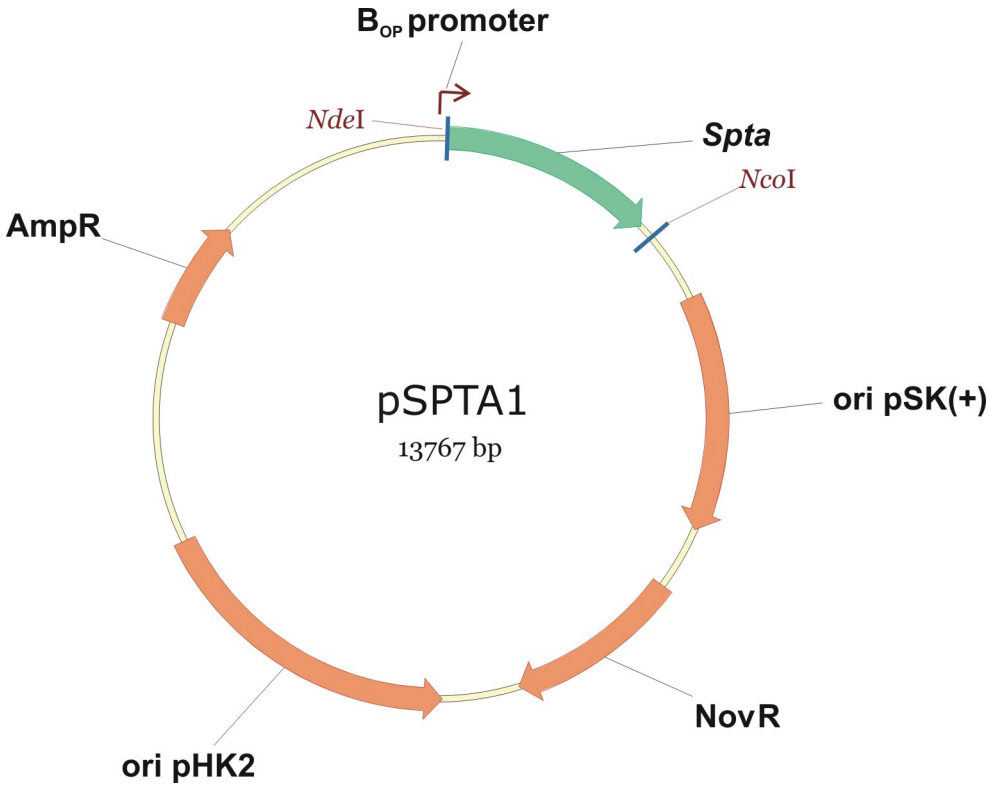

Figure 4.4 Vector pSPTA1. The vector pSTA1 was constructed by Shi *et al*., (2006) [172] by ligating the serine protease, SptA from *Natrinema* sp J7 into the *Nde*I/*Nco*I restriction sites of the vector pSY1 [130]. The AmpR gene encodes a *β-lactamase* providing ampicillin resistance and the NovR gene provides novobiocin resistance.

Site directed mutagenesis was performed on pSPTA1 using the primers shown in Table 4.2 to individually mutate the Tat motif's first and second arginine to a lysine, a twin lysine SptA mutant was also produced. The resulting plasmids were named pSPTAKR (R11K), pSPTARK (R12K) and pSPTAKK (R11K, R12K) which encode SptA-KR, SptA-RK and SptA-KK respectively.

	Forward Primer	Reverse Primer
SptA KK	CGTTATGTCCGGTGACAATA ACCAACACATGGATAAAAA ATCGCTTTTACAAACAGT	GCAATACAGGCCACTGTTATT GGTTGTGTACCTATTTTTTAG CGAAAATGTTTGTCA
SptA KR	CCGGTGACAATAACCAACACA TGGATAAAAGATCGCTTTTACA AACA	GGCCACTGTTATTGGTTGTGT ACCTATTTTCTAGCGAAAATG TTTGT
SptA RK	GTGACAATAACCAACACAT GGATCGAAAATCGCTTTTAC AAAC	CACTGTTATTGGTTGTGTACC TAGCTTTTAGCGAAAATGTTT G

Table 4.2 The primers used to mutate the twin arginines in the SptA twin arginine motif. Letters indicated in red are the amino acids that will be mutated by the primer. The SptA KR primers will mutate the first arginine to a lysine (R11K) while the primers SptA RK will mutate the second arginine to a lysine (R12K). The SptA KK primers mutate both arginines to create a twin lysine mutant (R11K, R12K). All primers are annotated 5'-3'.

H. volcanii H26 was then transformed with pSPTAKR, pSPTARK and pSPTAKK and the SptA plate assay was performed. The results are shown in Figure 4.5 which revealed that, as shown before by Shi *et al* (2006) [172], there was no secretion of the twin lysine mutant. Strikingly, also mutating the first or second arginine to a lysine blocked the translocation of SptA, showing that both arginines in the signal peptide of SptA are essential for translocation.

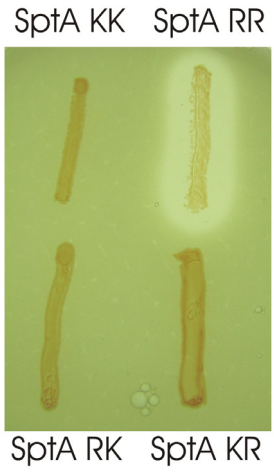

Figure 4.5 *H. volcanii* H26 transformed with pSPTAKR, pSPTARK and pSPTAKK. *H. volcanii* H26 was transformed with pSPTAKR, pSPTARK and pSPTAK and grown at 45ºC for 3 day on milk HV-YPC plate.

There are no anti-SptA antibodies available, so to allow for further biochemical studies PCR was used to incorporate a His tag onto the C-terminal of SptA. After amplification using the primers shown in Table 4.3 the gene was ligated into the *Nco*I/*Nde*I restriction sites of the plasmid pSY1 to form pSPTAHIS [172].

	Forward Primer	Reverse Primer
SptA C-terminal His Tag	ATAT**CATATG**TTTAGGAAG AATTTAATAGCGTGCT	ATAT**CCATGG**CTA**GTGATGG TGATGGTGATG**TCGACCGCG TTCGTC

Table 4.3 The primers used to attach a His Tag onto the C-terminal of SptA. The letters in blue indicate the *Nco*I/*Nde1* restriction sites and the underlined region is the sequence that encodes the 6 Histidine Tag. All primers are annotated 5'-3'.

H. volcanii H26 (pSPTAHIS) was grown on milk HV-YPC plates for 3 days at 45 ºC. Unfortunately, as revealed in Figure 4.6, the His tag appeared to significantly reduce the halo produced by SptA when compared to the positive control.

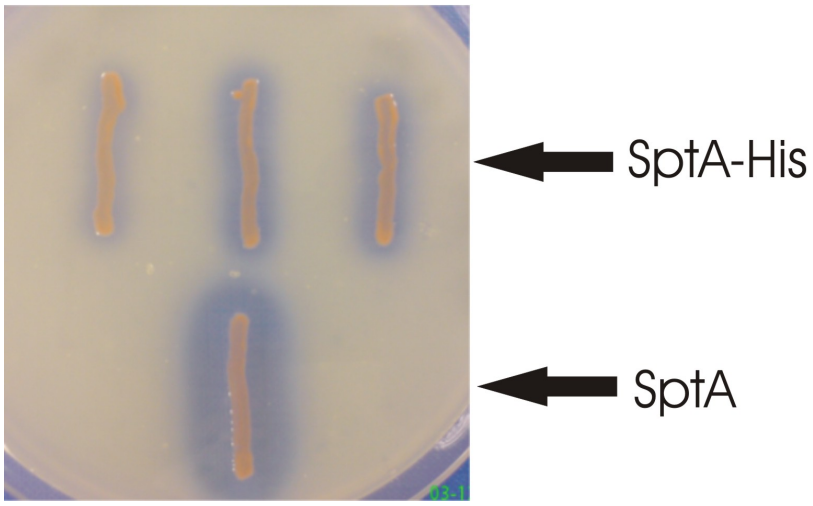

Figure 4.6 Secretion of SptA-His in *H. volcanii* H26. The top three colonies are separate transformations of *H. volcanii* H26 with pSPTAHIS. The bottom colony is the positive control of *H. volcanii* H26 (pSPTA). Colonies were allowed to grow for 3 days at 45°C on milk HV-YPC plate.

.

To quantify the reduction in proteolytic activity of strains expressing his-tagged SptA, four cultures of *H. volcanii* (pSPTAHIS) were grown to an OD_{600} of 0.8. After centrifugation the azocasein assay was performed on the supernatant and the results are shown in Table 4.4.

	Activity (%)
SptA-His	53 ± 5
SptA	100

Table 4.4 An azocasein assay performed on the supernatant of *H. volcanii* H26 (pSPTAHIS). Activity was calculated using the azocasein assay on the supernatant of *H. volcanii* H26 (pSPTA1) and 4 separate cultures of *H. volcanii* H26 (pSPTAHIS), background protease activity was subtracted and the average was calculated. The activity of *H. volcanii* H26 (pSPTA1) was taken to be 100 %. Calculations are shown in Appendix B.

Similar to the results in Figure 4.6, the azocasein assay indicates that the proteolytic activity is reduced by approximately 50% when the His tag is present. A Western blot was also performed, but the amount of SptA produced was apparently not sufficient to detect with anti-his antibodies (data not shown). These results made it difficult to use SptA for more detailed studies into its secretion by the Tat pathway.

4.3 The importance of the twin arginines for Tat translocation using the α-amylase AmyH

4.3.1 Introduction

A His tag onto the C-terminal of SptA greatly reduced its ability to translocate across the membrane (Chapter 4.2.). As antibodies are important for a detailed analysis of protein secretion, SptA was not very useful for further studies. Hutcheon *et al*., [150] identified and characterised an α-amylase secreted by *H. hispanica*, which was denoted AmyH. It is likely that this protein is also a Tat substrate since the AmyH signal peptide contains a clear twin arginine motif as shown in Figure 4.7.

Figure 4.7 The signal peptide of AmyH. The region in red shows the Tat motif and the region in blue shows the hydrophobic region of the signal peptide. The arrow indicates the processing site.

If it is possible to express AmyH in *H. volcanii* than this would allow for the analysis of its secretion as was done with SptA. If the amylase is secreted, it can degrade starch, which can be visualised on starch-agar plates that (after growth of colonies) are stained with iodine. A peptide tag would not be required as we already have anti-AmyH antibodies [134]. This plate assay would be used in conjunction with site directed mutagenesis to determine how specific mutations in the twin arginine motif affect the translocation of AmyH.

4.3.2 Developing an assay to confirm the translocation of AmyH

The primers in Table 4.5 were used to amplify the *AmyH* gene from the chromosomal DNA of *H. hispanica*. The gene was then ligated into the expression vector pSY1 by utilising the *Nde*I and *Nco*I restriction sites [172]. The resulting construct was named pAMY.

	Forward Primer	Reverse Primer
AmyH amplification	TTTGTTTAACTTTAAGAAGGAG ATATA**CATATG**AATCG	AAAA**CCATGG**GCTTTGTTAGCA GCCGGAT
AmyH KK Quick Change	CCGGCAGTAAGCAGGCGTCTAA GAAAACCGTTCTGAAAGGAATC G	GGCCGTCATTCGTCCGCAGATT CTTTTGGCAAGACTTTCCTTAG C

Table 4.5 Primers used to amplify AmyH and also to create an AmyH KK mutant. The underlined bases show the *Nco*I and *NdeI* restriction sites. The nucleotides that are shown in red are the amino acids where the mutations will occur. All primers are annotated 5'-3'.

H. volcanii H26 (pAMY) was plated onto starch HV-YPC plates. After three days incubation at 45°C the plates were stained with iodine solution and the results are shown in Figure 4.8. A halo formed around *H. volcanii* H26 (pAMY), but no halo was seen around the negative control (empty vector). This indicates that AmyH was capable of translocating across the *H. volcanii* membrane and consequently was able to break down the starch in the media.

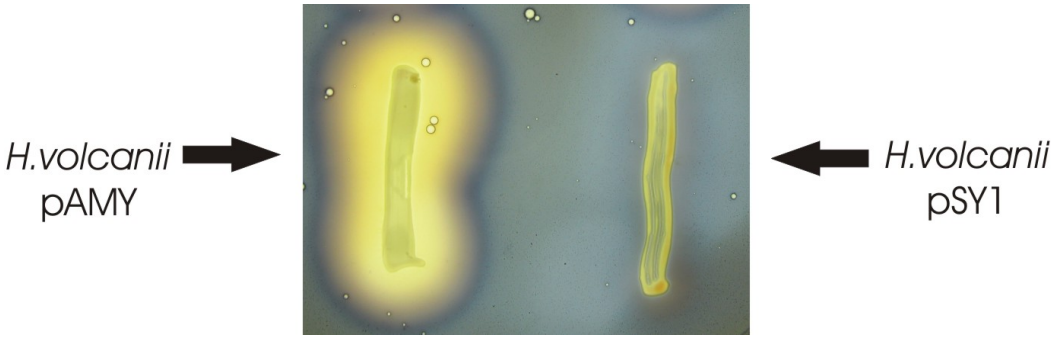

Figure 4.8 *H. volcanii* transformed with pAMY. *H.volcanii* (pAMY) and *H.volcanii* (pSY1) (negative control) were plated onto starch HV-YPC plates. After two days incubation at 45°C the plates were stained with iodine solution.

It was important to confirm that secretion of AmyH was indeed Tat-dependent. Therefore the twin arginines were mutated to twin lysines (R14K, R15K) by performing site directed mutagenesis, using the primers shown in Table 4.4. The vector encoding AmyH-KK (pAMYKK) was transformed into *H. volcanii* H26 and plated onto starch HV-YPC plates. After three days incubation at 45°C the plates were stained with iodine solution and the results are shown in Figure 4.9. Only a faint halo could be seen around the twin lysine mutant and this was most likely caused by cell lysis. Hence, this assay confirms that it was Tat-dependent translocation of AmyH that was responsible for the large halo produced by *H. volcanii* (pAMY), as it is unlikely that a Sec-dependent protein is affected by this twin lysine mutation.

Figure 4.9 *H. volcanii* transformed with pAMY and pAMYKK. Cultures of *H. volcanii* transformed with pAMY or pAMYKK were plated onto starch HV-YPC plates. After two day incubation at 45°C the plates were stained with iodine solution.

To compare the efficiency of AmyH production and secretion in *H. volcanii* with that of the native organism of AmyH, *H. hispanica* B3 (which is a transposon mutant that overproduces AmyH about 5-fold as compared to the wild type strain [150]) and *H. volcanii* (pAMY) were grown to an OD_{600} of 0.8. Cultures were then centrifuged and the starch-iodine assay was performed on both the cellular and the supernatant fractions. The results revealed that the *H. volcanii* supernatant (0.225 ± 0.018) had approximately 65% of the AmyH activity found in the *H. hispanica* B3 supernatant (0.342 ± 0.025), which was adequate for further experiments (Calculations are shown in Appendix B). No amylase activity was detected in the cellular fraction of *H. volcanii* (pAMY), indicating that it is the expression and not the secretion of AmyH that is less efficient in *H. volcanii* (data not shown).

4.3.3 The role of the twin-arginines in the AmyH signal peptide

The AmyH plate assay provided the means to determine that AmyH is a Tat dependent substrate. This makes it possible to investigate the importance of the two arginines found in the twin arginine motif. Site directed mutagenesis was performed to mutate the first arginine (R14K) or the second arginine (R15K) to a lysine. This was achieved by using the primers shown in Table 4.6. The vectors constructed were named pAMYKR and pAMYRK, encoding AmyH-KR and AmyH-RK respectively.

	Forward Primer	Reverse Primer
AmyH KR Quick Change	GGCAGTAAGCAGGCGTCTAA GCGAACCGTTCTGA	CCGTCATTCGTCCGCAGATTCG CTTGGCAAGACT
AmyH RK Quick Change	CAGTAAGCAGGCGTCTCGGA AAACCGTTCTGAAAGGAATC	GTCATTCGTCCGCAGAGCCTTT TGGCAAGACTTTCCTTAG

Table 4.6 The primers used to mutate each of the arginines in the AmyH twin arginine motif to lysines. The nucleotides that are indicated in red show where the lysine mutation will occur. All primers are annotated 5'-3'.

H. volcanii H26 (pAMYKR) and *H. volcanii* H26 (pAMYRK) were plated onto starch HV-YPC plates. After incubation for 3 days at 45°C the plates were stained with iodine solution. The results are shown in Figure 4.9, which revealed that transformants expressing the AmyH-KR or AmyH-RK mutants did not produce a halo. Interestingly, if *H. volcanii* H26 (pAMYKR) is incubated for 1 day at 45°C followed by 5 days at room temperature then a halo is produced that is approximately one third the size of the wild type amylase (Figure 4.10). No halo can be seen at room temperature for *H. volcanii* H26 transformed with pAMYRK or pAMYKK.

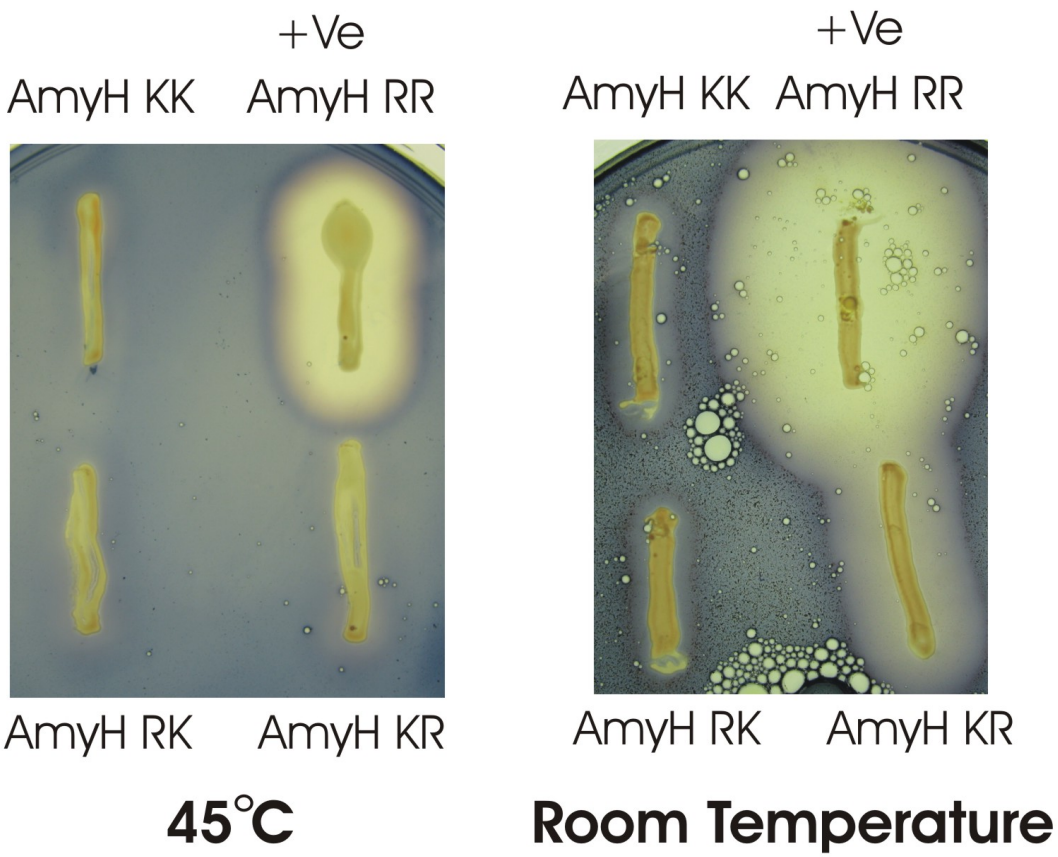

Figure 4.10 *H. volcanii* transformed with pAMY, pAMYKK, pAMYRK and pAMYKR. *H. volcanii* was transformed with pAMY (+ve), pAMYKK, pAMYRK and pAMYKR and allowed to grow on starch HV-YPC plates for 3 days at 45°C or 1 day at 45 °C followed by 5 days at room temperature. To visualise the halos the plates were stained with iodine solution.

To confirm the results from the 45°C AmyH plate assay, *H. volcanii* H26 transformed with pAMYRR, pAMYKK, pAMYKR and pAMYRK were grown for 6 days at 45°C. Cultures were centrifuged and the supernatant was concentrated 10-fold while the harvested cells were lysed by sonication. A Western blot was performed on both fractions and results are shown in Figure 4.11. AmyH is present in the cytoplasmic fraction of all transformants but as with the plate assay, there is only efficient translocation with the wildtype AmyH. Very small quantities of AmyH are detectible in the supernatant of the AmyH KK and AmyH KR mutants but none can be seen for the AmyH RK mutant. As expected, AmyH is present in the cellular and supernatant fractions of *H. hispanica* (positive control) but can not be detected in *H. volcanii* H26 (negative control).

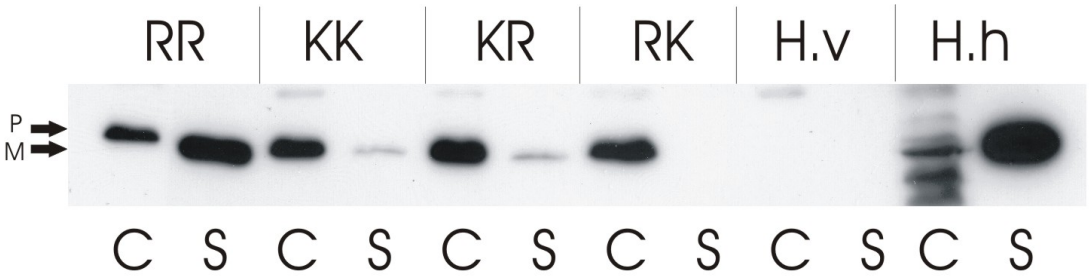

Figure 4.11 *H. volcanii* H26 expressing the AmyH twin lysine mutants. *H. volcanii* H26 was transformed with pAMY (RR), pAMYKK (KK), pAMYKR (KR) and pAMYRK (RK). After 6 days incubation at 45°C the transformants cellular (C) and supernatant (S) fractions were separated and a Western blot was performed. The same was done for the negative control, *H. volcanii* H26 (H.v) and the positive control *H. hispanica* (H.h). The arrow, P shows the precursor AmyH and the arrow, M shows the mature AmyH.

Pulse chase is a technique that provides information on the kinetics of protein translocation. The technique works by radiolabelling cells with ^{35}S-methionine (pulse) and then after 5 minutes adding an excess of "cold" methionine" (chase). At several time points during the chase, samples are taken and the protein of interest is purified by immuno-precipitation. By analysing these samples on SDS PAGE it possible to determine the rate at which the precursor protein is processed to mature protein and hence determine the kinetics of translocation. Pulse chase was performed on cultures of *H. hispanica* B3 (positive control) and *H. volcanii* H26 (pAMY). The results for *H. hispanica* B3 are shown in Figure 4.12 which revealed that the experiment was successful as you can clearly see the transition of precursor to mature AmyH in the positive control. Unfortunately, for reasons unknown, AmyH is not visible in *H. volcanii* during the 5 minute chase period (data not shown). As a consequence, pulse chase could not be used to investigate in more detail how the mutations of the AmyH signal peptide affect the rate of translocation.

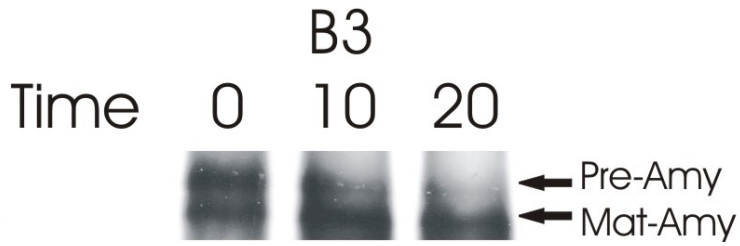

Figure 4.12 A pulse chase performed on AmyH expressed in *H. hispanica* B3. *H. hispanica* B3 was grown to an OD_{600} of 0.8 and then incubated in minimal media for a further hour. Pulse chase was performed and samples were removed at 0, 10 and 20 minutes after the chase. Immunoprecipitation separated AmyH which was then visualised on a SDS PAGE gel. Pre-Amy represents the precursor AmyH and Mat-Amy is the mature AmyH. For detection in *H. volcanii* H26 the film was exposed for 2 months but unfortunately bands were still not visible.

Since the AmyH-KR mutant gave the intriguing result of producing a halo when expressed in *H. volcanii* H26 at room temperature, it was of interest to see if the same was the case for SptA-KR mutant. So *H. volcanii* H26 transformed with pSPTAKR, pSPTARK and pSPTAKK were tested for SptA secretion at room temperature. Growth on HV-YPC milk plates revealed that none of the SptA mutants expressed in *H. volcanii* H26 produced a halo when the plates were incubated for 1 day at 45°C followed by 5 days at room temperature (Figure 4.13).

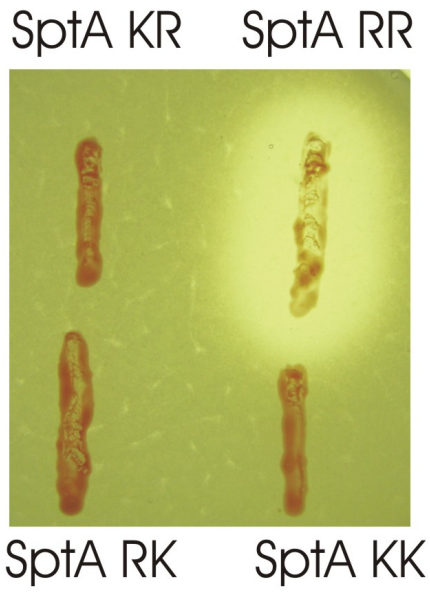

Figure 4.13 *H. volcanii* H26 transformed with pSPTAKR, pSPTARK, or pSPTAKK and grown at room temperature. Strains were grown on a milk HV-YPC plate for 1 days at 45° followed by 5 days at room temperature.

4.4 The importance of the other amino acids of the twin arginine motif for Tat translocation.

4.4.1 Introduction

The twin arginines found in the Tat signal motif have been found to be conserved in nearly all known Tat substrates. Experiments in Section 4.2 and 4.3 have confirmed the bioinformatic data that suggested the twin arginines also play an important role in *H. volcanii* for the translocation of proteins *via* the Tat pathway. In bacteria the twin arginine motif is (S/T)RRXFLK, with the five amino acids that surround the twin arginines occurring in a frequency greater then 50% [22]. Since these five residues are highly conserved it is likely they also play a role in Tat dependent translocation. This is, however, not the case for protein transport in chloroplasts where the twin arginines have been found to be the only conserved residues [173].

Stanley *et al*., (2000) [174] investigated the five surrounding amino acids in more detail using the *E. coli* Tat substrate SufI. The twin arginine motif of SufI is SRRQFIQ therefore the main difference from the classic motif is that a glutamine has replaced the lysine. There is also an isoleucine instead of the leucine although this has been found to be a common replacement [22]. Stanley *et al*., (2000) [174] mutated the first glutamine to a leucine and the other four amino acids to alanines to observe how these mutations affected translocation. A mutation of the final glutamine to a lysine was also made. It was found that mutating the serine and the first glutamine did not affect translocation while substitution of the isoleucine or the last glutamine to an alanine caused a slight decrease in translocation efficiently. Interestingly, when the phenylalanine was mutated to an alanine or the last glutamine to lysine then export was severely affected.

The bioinformatics in Section 4.1 indicate that Halophilic Tat substrates generally have a twin arginine motif that resembles (S/T)-R-R-X-(F/V/L)-L-(A/K). These data therefore suggests that the haloarchaeal Tat signal peptide is similar to the bacterial signal peptide, but with possibly a stronger preference for a leucine at position 6 and an alanine at position 7. The AmyH from *H. hispanica* has the motif SRRTVLK so

the main difference is that the valine has been replaced by a phenylalanine. It would be interesting to individually mutate these amino acids to alanines to see how it affects the translocation of AmyH and if these results relate to the bioinformatic analysis and the findings by Stanley *et al*., (2000) [174]. The reason for mutating the amino acids to an alanine is because this is a simple amino acid that is slightly hydrophobic; it does not affect the overall charge and normally has little effect on the structural properties of the signal peptide.

4.4.2 Mutation of the other 5 amino acids in the twin arginine motif and their effects on AmyH translocation

Site directed mutagenesis was performed on pAMY to mutate the codons of the five amino acids of AmyH that surround the twin arginines within the motif. Each amino acid was individually mutated to an alanine using the primers shown in Table 4.7 and the resulting plasmids were named pARRTVLK (S13A), pSRRAVLK (T16A), pARRTALK (V17A), pSRRTVAK (L18A) and pSRRTVLA (K19A).

	Forward Primer	Reverse Primer
AmyH Arrtvlk Quick Change	CAGTAAGCAGGCGGCTCGGCGA ACCGT	GTCATTCGTCCGCCGAGCCGC TTGGCA
AmyH srrAvlk Quick Change	GCAGGCGTCTCGGCGAGCCGTT CTGAAAG	CGTCCGCAGAGCCGCTCGGCA AGACTTTC
AmyH srrtAlk Quick Change	CGTCTCGGCGAACCGCTCTGAA AGGAATCGG	GCAGAGCCGCTTGGCGAGACT TTCCTTAGCC
AmyH srrtvAk Quick Change	GTCTCGGCGAACCGTTGCGAAA GGAATCGGAGTG	CAGAGCCGCTTGGCAACGCTT TCCTTAGCCTCAC
AmyH srrtvlA Quick Change	TCGGCGAACCGTTCTGGCAGGA ATCGGAGTGCTC	AGCCGCTTGGCAAGACCGTCC TTAGCCTCACGAG

Table 4.7 The Primers used to mutate the 5 amino acids that surround the twin arginines in the AmyH Tat motif. The nucleotides that are shown in red indicate where the mutation to an alanine occurs. The capital A shown in the primer name is the amino acid that the primer set will mutate. All primers are annotated 5'-3'.

H. volcanii H26 was transformed with the resulting five plasmids and plated onto starch HV-YPC plates. After 3 days incubation at 45°C the colonies were stained with iodine solution. The results are shown in Figure 4.14 which revealed that mutating the serine, threonine or lysine to an alanine has little effect on the size of the halo compared to the positive control. However, mutation of the valine or the leucine to an alanine appears to completely block the translocation of AmyH.

Figure 4.14 *H. volcanii* H26 transformed with the mutated pAMY vectors. Site directed mutagenesis was performed to individually mutate the 5 amino acids surrounding the twin arginines of AmyH. These vectors were then transformed into *H. volcanii* H26 and colonies were plated onto starch HV-YPC plates. After 3 days incubation at 45°C plates were stained with iodine solution. The amino acid marked as A indicates the position at which the amino acid from the motif SRRTVLK has been mutated to an alanine. The positive control is *H. volcanii* transformed with pAMY.

To look in more detail at the impact of these mutations, cultures of *H. volcanii* H26 transformed with the mutated pAMY vectors were grown for 6 days at 45°C. Cells were harvested by centrifugation and then the supernatant was concentrated 10-fold while the cells were lysed by sonification in water. A Western blot was performed on both fractions and the results are shown in Figure 4.15. As expected, the blot produced similar results to the plate assay. AmyH is present in the cytoplasm of all cultures but is not translocated if the valine or the leucine are mutated to an alanine.

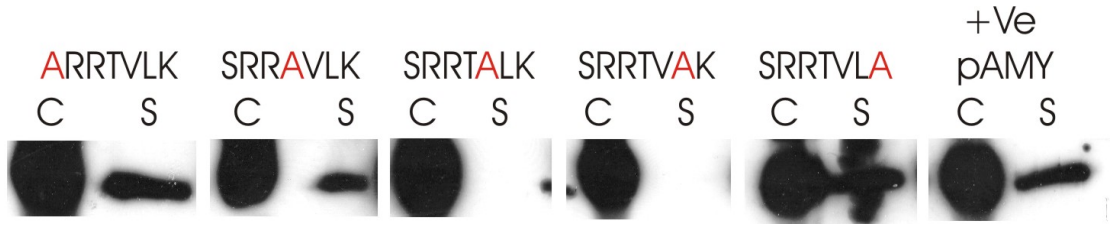

Figure 4.15 *H. volcanii* H26 transformed with the mutated pAMY vectors. The mutated vectors were transformed into *H. volcanii* H26 and 50 ml cultures were grown for 6 days at 45°C. The cellular (C) and supernatant (S) fractions were then resolved on by SDS PAGE and a Western blot was performed. The amino acid marked as A indicates the position at which the amino acid from the motif SRRTVLK has been mutated to an alanine. The positive control is *H. volcanii* transformed with pAMY.

It was a surprising observation that by growing *H. volcanii* (pAMYKR) at room temperature produced a halo that was not seen at 45°C. Thus, it is important to also

determine the effect of expressing AmyH at room temperature when it contains these five single base mutations. Therefore, colonies of *H. volcanii* containing these vectors were plated onto starch HV-YPC plates and incubated at 45ºC for one day, plates were then transferred to room temperature and allowed to grow for a further 5 days. As expected, mutating the serine (S13A), threonine (T16A) or lysine (K19A) to an alanine produced a halo as is also seen when grown at 45°C (See Figure 4.16). The mutation of the valine (V17A) or the leucine (L18A) to an alanine does not produce a clearing at 45°C. At room temperature, however, a halo approximately three quarters of the size of the other 3 mutations was formed by the valine mutation (V17A). Again, no significant starch degradation is seen by the expression of the leucine mutation at room temperature (L18A).

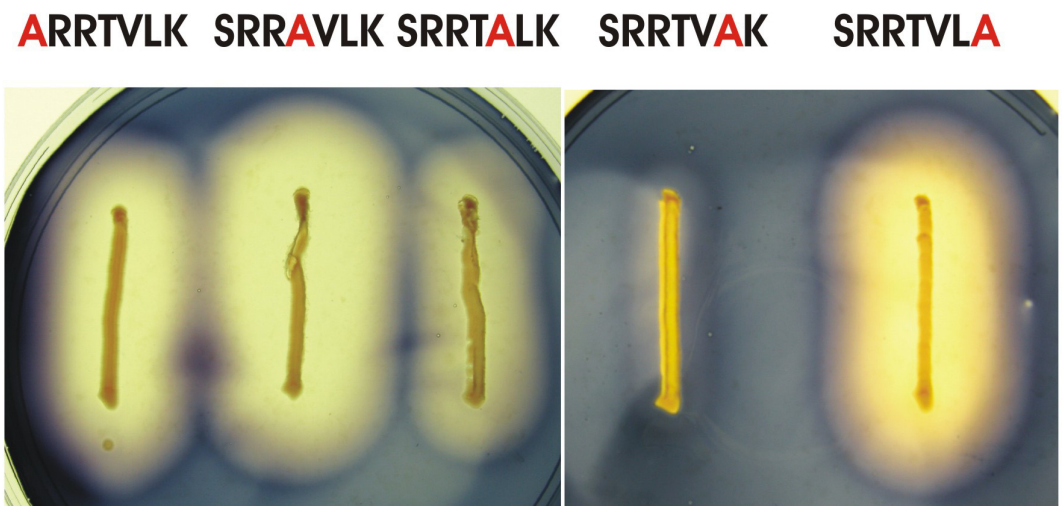

Figure 4.16 *H. volcanii* H26 transformed with the mutated pAMY vectors and grown at room temperature. Site directed mutagenesis was performed to individually mutate the 5 amino acids surrounding the twin arginines of AmyH. These vectors were then transformed into *H. volcanii* and colonies were plated onto starch HV-YPC plates. After 1 day incubation at 45°C and then 5 days at room temperature the plates were stained with iodine solution. The amino acid marked as A indicates the position at which the amino acid from the motif SRRTVLK has been mutated to an alanine.

4.5 The use of an *in vitro* translocation assay as a method to investigate the translocation of AmyH

4.5.1 Introduction

We have developed an *in vitro* assay to investigate the bioenergetics involved in the translocation of AmyH. The basic principle is to synthesize radiolabelled preAmyH and import this into inverted membrane vesicles (IMVs), which are sealed membranes with an orientation opposite to that of intact cells [152]. With isolated IMVs it is possible to control the salt concentration and composition, both internally and externally. In native conditions in a haloarchaea, the cytoplasm has a high concentration of KCl, while the external medium has a high concentration of NaCl. This is mimicked in IMVs using a high external KCl and a high internal NaCl concentration [175]. Dr. Bolhuis has already shown that AmyH could be successfully translated and is capable of translocating into IMVs. There is also evidence indicating that processing of the signal peptide occurs (see Figure 4.17) [152].

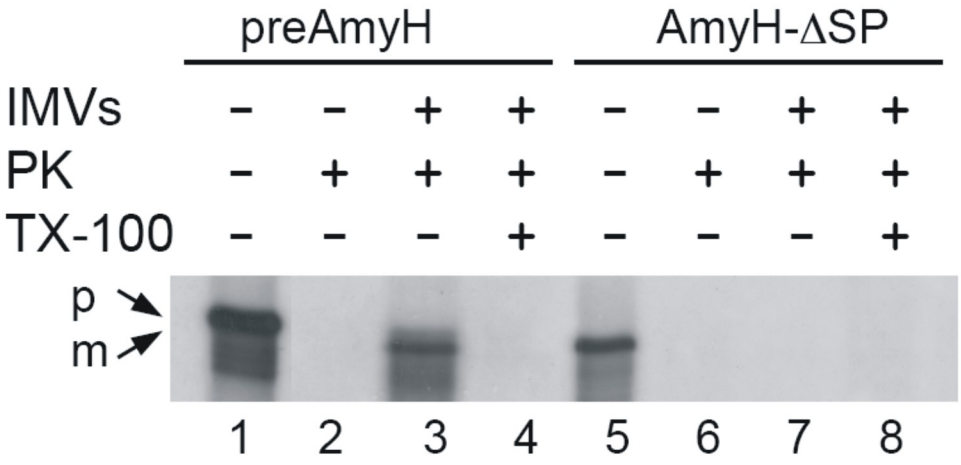

Fig. 4.17. *In vitro* translocation assay of preAmyH. Lanes 1–4 and 5–8 show preAmyH and AmyH-ΔSP, respectively. PreAmyH and AmyH-ΔSP were incubated in the presence or absence of IMVs, proteinase K (PK) and/or Triton X-100 (TX-100) as indicated. The loading for the translation reactions in lanes 1 and 5 is 5% of the amount used in the translocation assays in lanes 2–4 and 6–8. This image was taken from Kwan *et al*., (2008) [152]. The arrow, P shows the precursor AmyH and the arrow, M shows the mature AmyH.

It was important to confirm that these results were a direct result of Tat–dependent translocation of AmyH. This is possible by performing the *in vitro* translocation assay (Figure 4.18) on AmyH and an AmyH-KK (R14K, R15K) mutant. AmyH and its twin lysine mutant are labelled with [^{35}S]-methionine and allowed to translocate into energised IMVs. Proteinase K is then added to degrade any protein that was unable to translocate. If AmyH is indeed a Tat substrate then it will be protected by the IMVs, while the twin lysine mutant is not imported and therefore is degraded.

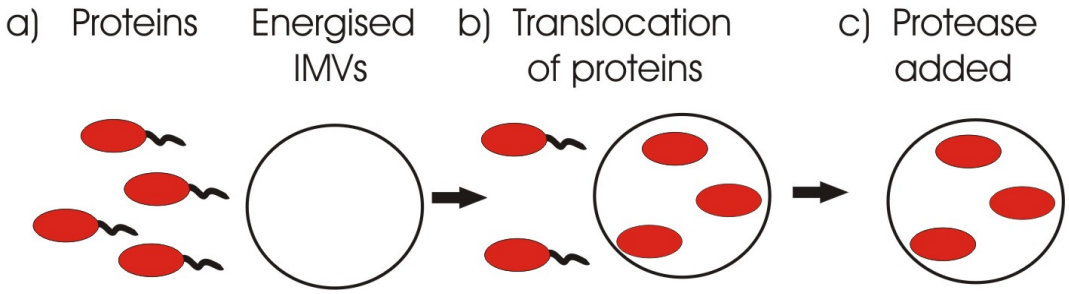

Figure 4.18 The *in vitro* translocation assay involving inverted membrane vesicles. a) the proteins are added to the energised IMVs. b) the Tat substrates translocate into the IMVs while proteins unable to utilise the Tat pathway and thus cannot translocate will remain on the outside. c) proteinase K is added which degrades the proteins that have failed to translocate while the Tat proteins are protected by the IMVs.

4.5.2 The use of an *in vitro* translocation assay to determine if AmyH is a Tat substrate

Dr A. Bolhuis kindly provided the plasmid pET-AMYH shown in Figure 4.19. It was constructed from the plasmid pET21a with the *H. hispanica AmyH* gene ligated behind its T_7 promoter.

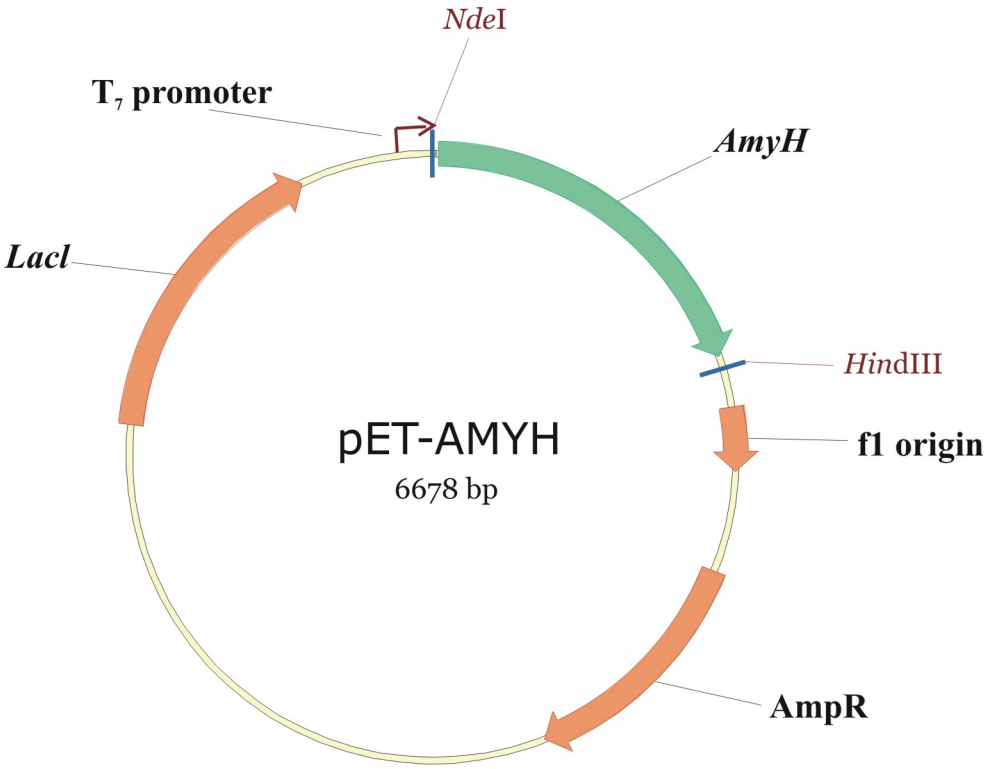

Figure 4.19 The vector pET-AMYH. The *H. hispanica* AmyH gene was ligated in front of the T_7 promoter of the pET21a vector by utilising the *Nde*I/*Hin*dIII restriction sites. The plasmid also encodes a *lacI* gene which is a repressor gene and a gene encoding β-lactamase (AmpR) which provides ampicillin resistance.

Site directed mutagenesis was performed to mutate the twin arginines in the AmyH Tat motif to twin lysines (R14K, R15K) using the primers shown in Table 4.5. The resulting construct was named pET-AMYKK.

In vitro translation was carried out on pET-AMYH and pET-AMYHKK with the resulting labelled AmyH and AmyH-KK being run on a SDS PAGE gel. ImageJ (http://rsbweb.nih.gov/ij/) was used to calculate the ratio between these two proteins so that similar quantities could be used in future experiments (data not shown). It was also important to confirm that the translated AmyH was capable of folding when added to a high salt buffer and did not, for example, require cytoplasmic chaperones as was suggested by Hutcheon *et al*., (2005) [150]. Therefore the susceptibility of AmyH to protease degradation was tested in the presence/absence of 2.5 M KCl. If AmyH is capable of folding in the high salt buffer then it will be less susceptible to protease degradation compared to the unfolded AmyH in buffer lacking salt. The translated AmyH was allowed to fold in buffer (50 mM Tris-HCl pH 7.5, 1 mM $CaCl_2$, 25 mM $MgCl_2$ and 5 mM DTT +/- 2.5 M KCl) for 10 minutes. Proteinase K (0.5 mg/ml) was added and then incubated at 37°C for 10, 30 and 60 minutes. Samples were removed and placed on ice for 30 minutes in the presence of the 1 mM PMSF (inhibits proteinase K). Finally, samples were run on a SDS PAGE gel and results are shown in Figure 4.20.

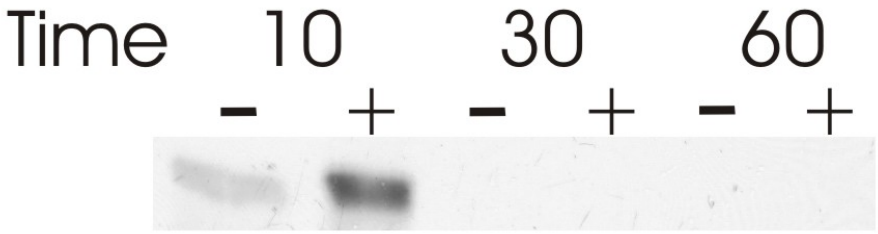

Figure 4.20 Susceptibility of translated AmyH to protease digestion with and without the presence of 2.5 M KCl. Translated AmyH was added to Buffer +/- 2.5 M KCl and incubated for 10, 30 and 60 minutes. Protease digestion was then inhibited and samples run on a SDS PAGE gel. The Time is shown in minutes. - = Buffer lacking 2.5 M KCl and + = Buffer containing 2.5 M KCL.

Analysis of the results using ImageJ (http://rsbweb.nih.gov/ij/) show that after 10 minutes there is four times more translated AmyH in the presence of KCl compared to AmyH in the absence of salt. It is important to note that proteinase K still has about 70% of its activity in 2.5 M KCl as compared to its activity in the absence of KCl (As shown by "Hutcheon, (2008)) [176], however this would not account for the 4-fold difference in activity. Thus this result shows that AmyH translated *in vitro* is less susceptible to protease digestion when added to a high salt buffer. This means

that AmyH has formed a more stable structure in these conditions, which is indicative of proper folding. These results are also corroborated by data of Hutcheon *et al.*, (2005) [150], who found that AmyH remains active in low salt concentrations but has taken up a more loosely folded confirmation. After 30 minutes AmyH is no longer detectable.

Next, inverted membrane vesicles were produced and *in vitro* translocation was performed on equal amounts of the labelled AmyH and AmyH-KK. After TCA precipitation the samples were run on a SDS PAGE gel and the results are shown in Figure 4.21. Wild type AmyH could be detected when the IMVs were placed in detergent which confirms that the amylase translocated successfully. However, the AmyH-KK mutant was barely detectible, indicating that it was degraded by proteinase K, providing evidence that it was indeed unable to translocate *via* the Tat pathway *in vitro*.

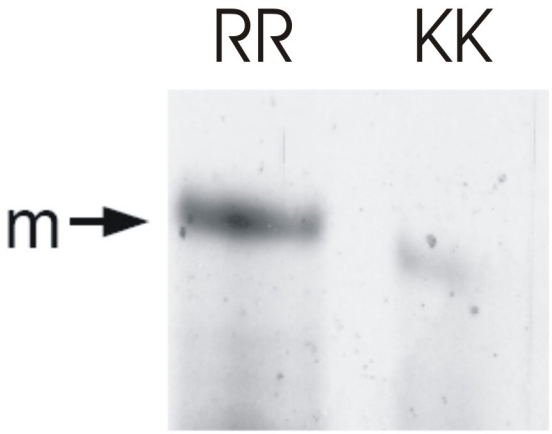

Figure 4.21 *In vitro* translocation of AmyH and AmyH KK. *In vitro* translocation was performed on equal amounts of the translated AmyH and AmyH KK. After protease digestion and TCA precipitation both samples were run on a SDS PAGE gel. The RR lane represents the *in vitro* translocation of wildtype AmyH while the lane KK shows the *in vitro* translocation of the twin lysine AmyH mutant.

4.6 The Halophilic Tat pathways ability to translocate inactive AmyH

4.6.1 Introduction

As mentioned previously, analysis of the known bacterial Tat substrates has shown that they are predominantly co-factor binding [169]. In fact, evidence suggests that if the insertion of the cofactor is prevented then the Tat substrate is no longer capable of translocating [22]. This would make sense as it prevents the protein from translocating *via* the Tat pathway prior to cofactor binding. Therefore, it seems logical to assume that there is some form of proof-reading system in place that ensures that the substrate is correctly folded. For example, Delisa *et al*., (2003) [19] attached a Tat signal peptide onto the *E. coli* alkaline phosphatase, PhoA. Normally PhoA translocates unfolded into the periplasm through the Sec system where the more oxidising environment allows the correct formation of disulphide bonds. It was possible to show that PhoA can only translocate *via* the Tat pathway in mutant strains that allow disulphide bond formation in the cytoplasm. Hence this suggests that some form of proof-reading system for correct folding does indeed exist. Meanwhile Richter *et al*., (2007) have suggested that there is an inbuilt tendency for the Tat pathway to only transport globular proteins that do not possess exposed hydrophobic patches and therefore again they must be correctly folded before translocation can occur.[177]

Hynds *et al*., (1998) [20] showed that the lumen protein, dihydrofolate reductase (DHFR) was capable of translocating across the chloroplasts thylakoid membrane *via* the ΔpH pathway if an appropriate signal peptide was attached to its N-terminal. However, when DHFR was truncated and therefore malfolded it did not affect its ability to translocate. The same was also seen when full length DHFR was malfolded by replacing the amino acid threonine with the analog β-hydroxynorvaline [20]. These results suggest that the proof reading system is linked to cofactor insertion and not reliant on the correct folding of the protein, or that proof reading only exists in the bacterial Tat system and not the thylakoidal system. Note that the thylakoid experiments were performed *in vitro* only, and these have not been confirmed *in vivo*.

Although AmyH does not bind a cofactor it would be interesting to see if the *H. volcanii* Tat pathway also possesses some form of proof reading system that would block translocation of mutant forms of AmyH. It was anticipated that a very mild mutation with respect to protein translocation would be one that affected activity of the protein but not its folding, thus in an effort to test this an active site mutant was produced. AmyH has three predicted active site residues with two aspartic acids residues and one glutamate residue as shown in Figure 4.22 [150]. It was anticipated that the mutation of one of these residues to an alanine would result in an enzyme that was inactive but still fully folded, since similar experiments showed this to be the case for other α-amylases, such as Amy1 from barley and Taka-amylaseA from A*spergillus oryzae* [178]. If this is true for AmyH then it would then be possible to test whether this inactive AmyH is still capable of translocating *via* the Tat pathway.

```
A       1              2              3              4
HHI  ADAVINH 137  GIRWDAAKH 221  WTVGEVLD 248  FVSNHD 308
NAM  VDIVLNH 167  GLRIDAAAH 251  WRVGEVWD 278  FVQNHD 340
PWO  ADVVINH 149  GWRFDYVKG 240  WAVGEYWD 263  FVANHD 327
THY  ADIVINH 133  AWRFDYVKG 224  WAVGEYWD 247  FVANHD 301
       * *:**         * *          * ***  *      ** ***
```

Figure 4.22 The active sites of a variety of α-amylases. The amylases that are aligned are from *H. hispanica* (HHI), *N. amylolyticus* (NAM), *Pyrococcus woesei* (PWO) and *Thermococcus hydrothermalis* (THY). The active site residues are shown in bold. * are identical residues and : are conserved residues [150].

4.6.2 The ability of inactive AmyH to translocate via the Tat pathway

The aspartic acid which forms the first part of the active site of AmyH (AS1) was mutated to an alanine (D217A) using site directed mutagenesis. The primers that were used are shown in Table 4.8, and the resulting vector was named pAMYAS1.

	Forward Primer	Reverse Primer
AmyH AS.1 Quick Change	GGT ATT CGC TGG GCC GCC GCC AAA CAC	CCA TAA GCG ACC CGG CGG CGG TTT GTG

Table 4.8 The primers used to mutate the active site 1 of AmyH from an aspartic acid to an alanine. The nucleotides that are shown in red indicate where the mutation to an alanine occurs. Both primers are annotated 5'-3'.

H. volcanii H26 (pAMYAS1) was grown at 45°C for 3 days on starch HV-YPC plates. Colonies were stained with iodine solution and the results are shown in Figure 4.23a. A halo can clearly be seen for the positive control but not for the AS.1 mutant. It is impossible to determine from the plate assay if the mutation has prevented AmyH from translocating as it is no longer in an active conformation. Therefore, *H. volcanii* H26 (pAMYAS1) was grown for 6 days at 45°C. After centrifugation the supernatant was concentrated 10 fold while the harvested cells were lysed by sonification in water. A Western blot was performed on the separate fractions and the results are shown in Figure 4.23b. The blot revealed that the mutation to the AmyH active site is preventing its translocation as amylase was only detectable in the cellular fraction.

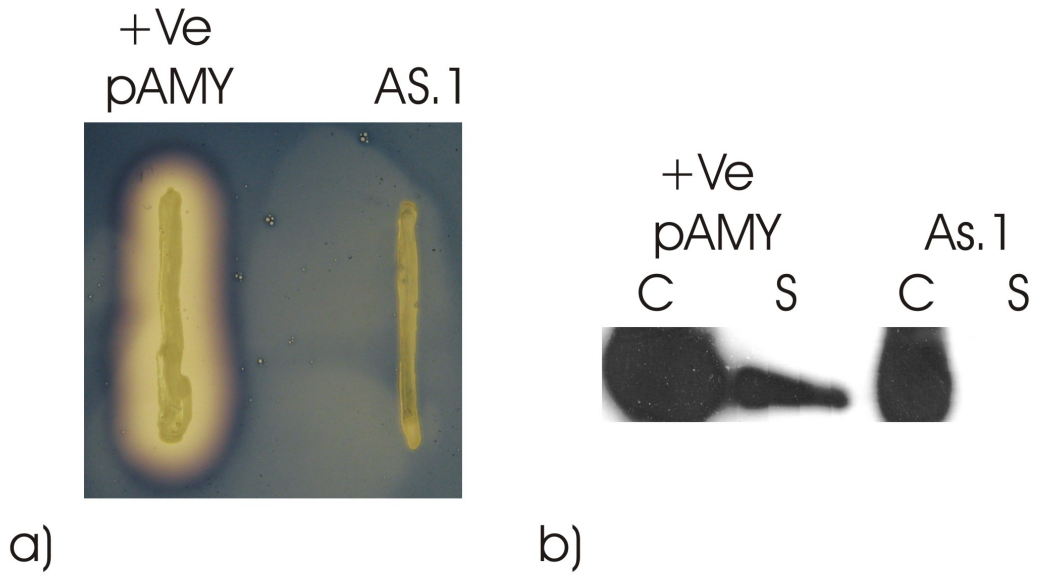

Figure 4.23 The effect of the AS.1 mutation on the translocation of AmyH A) pAMY (positive control) and pAMYAS1 (AS.1) were transformed into *H. volcanii* H26 and plated onto starch HV-YPC media. After incubation for 3 days at 45°C the plates were stained with iodine solution. B) pAMY (positive control) and pAMYAS1 (AS.1) were transformed into *H. volcanii* H26 and incubated for 6 days at 45°C. The cellular (C) and the supernatant (S) were then separated and run on a SDS PAGE gel and a Western blot was performed.

To test the folding of AmyH-AS1 in more detail, it was attempted to purify the protein on a β-cyclodextrin column, as binding to this column would indicate that the protein was correctly folded. To this purpose, *H. volcanii* H26 (pAMY) and *H. volcanii* H26 (pAMYAS1) were grown for 5 days at 45ºC and then cells were harvested by centrifugation. Cells were lysed by sonification in water and then small scale purification was performed using a β-cyclodextrin-sepharose column. The resulting fractions were run on a SDS PAGE gel and a Western blot was performed. The results are shown in Figure 4.24 which revealed that the correctly folded AmyH binds efficiently to the column but the mutation in AS.1 prevented the amylase from binding.

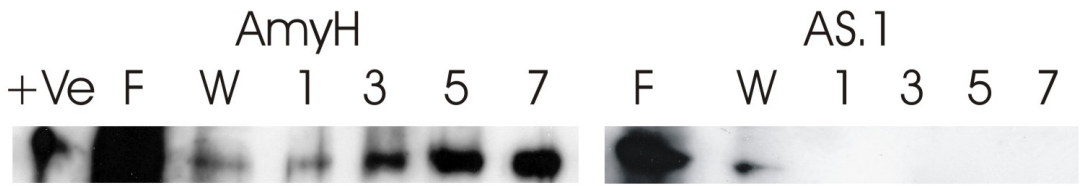

Figure 4.24 Purification of AmyH and AmyH-AS.1. 50ml cultures of *H. volcanii* H26 transformed with pAMY (AmyH) and pAMYAS1 (AS.1) were grown for 6 days at 45°C. Purification was then performed using a β-cyclodextrin-sepharose column. The resulting fractions were run on a SDS PAGE gel and a Western blot was performed. The +ve control was a sample of previously purified AmyH. F = the flow through and the W = the wash fraction. 1, 3, 5 and 7 shows the corresponding elution fractions.

It is important to note that the mutated AmyH (D217A) not being able to bind to the column does not confirm that the amylase is unfolded since the active site and the binding site may occur at the same position [179]. Thus it could be that the mutation of the aspartic acid (D217A) also disrupts the binding site preventing AmyH from binding to the β-cyclodextrin on the column. Therefore, further experiments are required to determine the amylases conformation. A protease-susceptibility assay would reveal if the active site mutation causes the protein to become unfolded. Results can be compared to a twin lysine AmyH mutant that is also incapable of translocating but is likely to have remained in a folded confirmation. Thus *H. volcanii* H26 (pAMYKK) and *H. volcanii* H26 (pAMYAS1) were grown for 3 days at 45°C. The cells were harvested and lysed by sonification in water (the addition of water made a final concentration of ~ 1 M salt). A protease assay was then performed by adding protease K (0.25mg/ml) to the cellular fractions and incubating for 5, 10 and 20 minutes at 37°C. Samples were immediately placed on dry ice and 1 mM PMSF was added to inhibit protease K. After 30 minutes the samples were run on a SDS PAGE gel and a Western blot was performed. The results are shown in Figure 4.25 which revealed that the AS.1 mutation caused AmyH to become more susceptible to protease digestion. This suggests that the mutation has led to an incorrectly or partially folded enzyme.

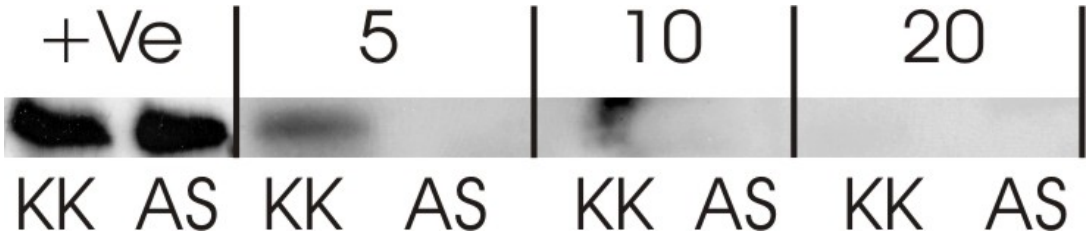

Figure 4.25 Protease sensitivity of the AmyH-KK and AmyH-AS.1 mutants. *H. volcanii* H26 transformed with pAMYKK (KK) and pAMYAS1 (AS) were grown for 3 days at 45°C and then a protease assay was performed on the cytoplasmic fraction. The positive control is the cytoplasmic fractions prior to protease digestion so that equal quantities of the proteins could be used in the assay. 5, 10 and 20 show the time (minutes) that the samples were treated with protease K.

4.7 Discussion

The classic bacterial Tat motif has been designated as (S/T)-R-R-x-F-L-K [35], but it was important to determine if the haloarchaeal Tat proteins also utilise this motif. The *H. hispanica* Tat substrate AmyH focussed on in this Chapter has the motif SRRTVLK which suggest that this is indeed the case. The bioinformatic analysis of *H.volcanii*, *H. marismortui, Halobactrium* sp NRC-1 and the haloalkaliphile, *N. pharaonis* confirms this with the majority of proteins containing the motif (S/T)-R-R- x-(F/V/L)-L-(A/K). The main result of interest is the possible importance of the leucine at position 6, which is more dominant in halophilic organism than in *E. coli*. The other result of note is the preference towards alanine at position 7 of the halophilic Tat substrates which is not found to be the case with *E.coli*; TorA is the only Tat substrate that contains an alanine at this position [23].

As mentioned previously, it has also been shown that by replacing the H-domain of the *E. coli* Tat substrate TorA with a more hydrophobic region resulted in the protein to be rerouted *via* the Sec pathway [32]. It was therefore important to determine if halophiles also have this noticeable distinction in the hydrophobicity of Sec and Tat substrates. Results revealed that this is indeed the case with the Sec substrates from *Halobacterium* sp NRC-1 being more hydrophobic than their Tat counterparts, similar to what was found for *E.coli*. Therefore this suggests that the H-domain may be another decisive factor for directing halophilic proteins to their respective pathways.

Shi *et al*., (2006) [172] have already shown that the SptA plate assay provides a useful tool for confirming successful translocation *via* the Tat pathway. Section 4.2 utilised this technique to show that SptA translocation is blocked when either the first or second arginine of the Tat motif is mutated to a lysine. Ideally, a Western blot could also be performed to confirm these results as it a much more sensitive technique then the plate assay. Unfortunately, attaching a His tag onto the C-terminal of the protease severely affected its translocation. This is possibly due to the increase in positively charged residues from the His tag, causing SptA to become stuck to the negatively charged membrane surface during translocation. Although it cannot be assumed that

it is the translocation of SptA that is affected as, for example, the His-Tag could cause a reduction in expression or affect the stability of its mRNA. Therefore, ideally the amount of SptA in the cytoplasm would be measured, but there are several cytoplasmic proteases that result in a high background of proteolytic activity, while the lack of antibodies against SptA means that a Western blot can not be performed.

The availability of anti-AmyH antibodies meant that a detailed investigation into the twin arginine motif of AmyH could be performed. Plate assays showed that there was minimal translocation when the twin arginines were individually mutated to lysines, similar to the results obtained with the SptA protein. Analysis of medium and cellular fractions by Western blotting indicated that there was some translocation occurring with the KK and KR mutants albeit at a greatly reduced rate. However, these results need to be confirmed in case of inconsistencies caused by the concentration step, for example AmyH becoming stuck to the membrane. No translocation could be seen for the RK mutant. The importance of these twin arginines appear to vary greatly depending on the specific protein. For example, individually mutating either of the arginines to a lysine can be tolerated in the *E.coli* Tat substrate, SufI although translocation is greatly reduced [174]. Ize *et al*., (2002) [37] has even shown that Colicin V with a Tat signal peptide containing an RR->KK mutation was still translocated, albeit at greatly reduced rate. However, in other examples the mutation of just one arginine completely abolishes translocation, as seen with for instance the multicopper oxidase YacK from *E.coli*, or the nitrous oxide reductase from *Pseudomonas stutzeri* [180].

When *H. volcanii* H26 (pAMYKR) was allowed to grow at room temperature, some apparent secretion was observed with a halo approximately one third of the size of that produced by *H. volcanii* H26 (pAMY). A similar effect was seen with the V17A mutant (see also below). One possible explanation is that the AmyH R14K and V17A mutants become stuck in the translocase, which in turn could lead to increased cell lysis. Ize *et al*., (2003) [24] have already shown that if the *E. coli* Tat pathway is blocked then cells become much more sensitive to SDS. Unfortunately, SDS precipitates in high salt concentrations so a similar technique could not be used on *H. volcanii* expressing AmyH-KR or AmyH-KK. An alternative explanation is that at lower temperatures *H. volcanii* H26 becomes more tolerant to the translocation of

Tat substrates with specific mutations in their signal peptide. Interestingly, a KR mutation in the signal peptide of SptA did not lead to a release of SptA at room temperature in *H. volcanii* H26. This suggests that there is possibly a difference between the interactions of the signal peptides of AmyH and SptA with the *H. volcanii* Tat system. An experiment that could provide better insight into this would be to swap the wild-type and mutant signal peptides of AmyH and SptA, and see how these mutations affect protein secretion at 45 ^{o}C and at room temperature.

The importance of these amino acids surrounding the twin arginines was analysed by using mutagenesis and the results were compared to similar experiments performed on the *E. coli* Tat substrate, SufI [174]. Results revealed that a mutation of the serine, threonine or lysine to an alanine had no effect on the translocation of AmyH. The same was found when mutating the serine and threonine of SufI to an alanine. However, SufI naturally has a glutamine instead of the consensus lysine and, interestingly, mutation of this amino acid to a lysine blocks translocation [174]. However, further experiments by Stanley *et al*., (2000) [174] showed that similar to AmyH, mutating the naturally occurring lysine of the *E. coli* Tat substrate, YacK to an alanine also had no effect on translocation.

Mutation of the valine or the leucine in the AmyH signal peptide to an alanine blocked translocation completely. It is unusual for valine to occupy position 5 of the motif as the phenylalanine found in the bacterial motif, after the invariant arginines is the most highly conserved amino acid (80%) [174]. Stanley *et al*., (2000) [174] found that mutating the phenylalanine to an alanine completely blocked translocation while a mutation to a leucine had no effect. They suggested that a very hydrophobic residue is essential at this position. This would explain the presence of the hydrophobic valine in AmyH and why replacement to an alanine blocked translocation. Interestingly, mutation of the leucine in the SufI Tat motif to alanine only showed a small decrease in the rate of export yet in AmyH this mutation blocked transport completely. Berks (1996) [22] bioinformatically analysed the bacterial twin arginine motif and suggested that, although this leucine is highly prevalent, it is not essential for translocation. Therefore, to my knowledge the importance of this leucine in Tat translocation appears to have only been shown with AmyH; thus, it is possible that the importance of this leucine is specific to halophilic Tat substrates. This

corroborates the earlier bioinformatic data which had already given an indication into the importance of this leucine for the translocation of halophilic Tat substrates.

Chapter 4 also aimed to investigative the possibility that the Tat complex of *H. volcanii* can translocate malfolded proteins. It has already been shown that this is possible for bacteria with co-factor binding being more essential then correct folding [20, 22]. It was of interest to look at the halophilic Tat pathway since such a large proportion of their secretory proteins utilise this system. Even though the majority of these proteins are not predicted to bind co-factors, the system may require more stringent proof-reading since the pathway plays such an important role to the organism. The first active site of AmyH was mutated to an alanine which in turn prevented its secretion. Purification of the amylase was not possible as the mutation to AmyH also prevented its binding to a β-cyclodextrin column. Separate binding and catalytic sites have been found in approximately 10% of known amylases, but alignments comparing AmyH to the amylases investigated by Machovič and Janeček (2006) confirmed that AmyH is not one of these [179]. Hence it is likely that this mutation of the active site prevented AmyH from binding to the column. This is in agreement to findings by Takase (1994) [181] who individually mutated all three active site mutations of the *B. subtilis* α-amylase (BSUA), which led to a loss in both catalytic activity and the ability to bind its substrate. BSUA shares 23% similarity to AmyH and also does not contain a separate binding site. The active site mutation caused AmyH to become more susceptible to degradation, which indicated that the protein no longer folds correctly. The active site mutant was also not secreted, which would suggest that the Tat system of *H. volcanii* indeed has a proof reading system, but further experiments are required to confirm this.

To conclude, both the SptA and the AmyH assays revealed that both of the arginine residues in the Tat motif are critical to translocation. Interestingly, the AmyH-KR (R14K) and the AmyH valine (V17A) mutants when expressed at room temperature either cause secondary effects leading to cell lysis or the Tat pathway is more susceptible to their translocation under these conditions. The AmyH assay was also used for studies into the other five amino acids of the Tat motif. Results mainly corresponded to the findings of the *E. coli* Tat substrate SufI, except for the mutation of the leucine (L18A) which surprisingly blocked translocation. Finally, the ability of

the *H. volcanii* Tat pathway to translocate an inactive protein was investigated. Unexpectedly, the mutation in the first active site of AmyH affected the amylases ability to fold which in turn prevented its secretion. This suggests the possibility that a proof reading system for the Tat pathway in *H. volcanii* does indeed exist.

Chapter 5: Bioenergetics of Haloarchaea

5.1 The Bioenergetics of the Halophilic Tat pathway

5.1.1 Introduction

Yahr and Wickner (2001) [45] have shown that the *E. coli* Tat pathway is driven solely by the PMF. More specifically, the Tat substrate pre-SufI requires only the ΔΨ and not the ΔpH for translocation [182]. Meanwhile, the equivalent thylakoid ΔpH dependent pathway is mainly driven by the ΔpH although the ΔΨ does have some minor involvement [183, 184]. Interestingly, in haloarchaea there is a strong sodium gradient available which provides another potential source of energy and it has been suggested that it is the SMF rather then the PMF that drives many haloarchaeal processes [152].

Ionophores are often used to investigate the bioenergetics of cellular processes in prokaryotes since they disrupt various gradients across the membrane. For example, carbonyl cyanide m-chlorophenylhydrazone (CCCP) is a proton carrier and uncoupler that disrupts the entire proton motive force, while the ionophore monensin is an antiporter with a high affinity for Na^+ ions that dissipates the chemical gradient of the SMF (Figure 5.1) [169, 185].

Three other ionophores of interest when looking at the bioenergetics of an organism are valinomycin, nonactin and nigericin, which all disrupt the membrane gradients in different ways from CCCP or monensin [186]. Valinomycin is a K^+ specific ionophore that dissipates the ΔΨ (Figure 5.2). Nonactin functions in a similar way with high affinity to K^+ ions but there is also evidence to suggest that there is some affinity to Na^+ and NH_4^+ ions. Finally, nigericin is an antiporter like monensin but it has a higher specificity to K^+ ions so will dissipate the chemical gradient of the PMF rather then the SMF [152].

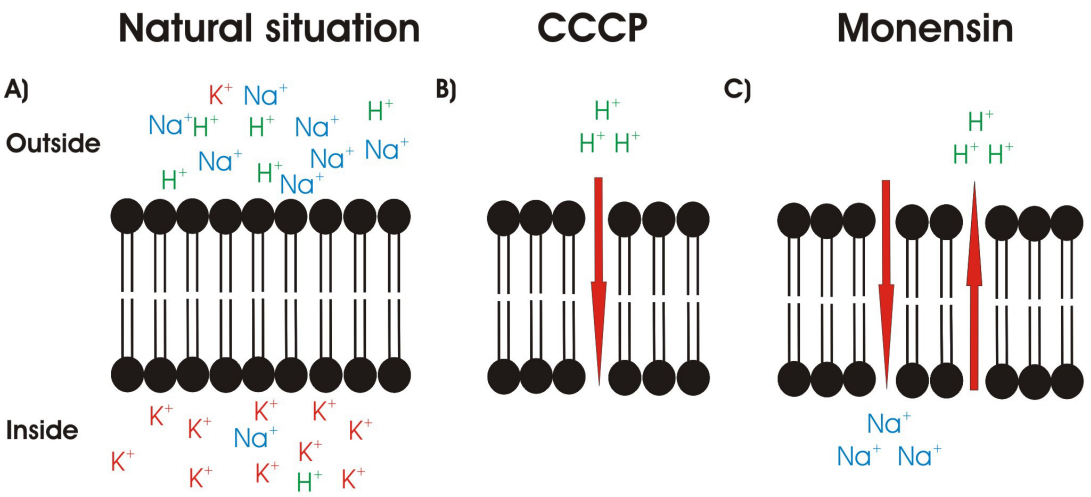

Figure 5.1 The effect of the ionophores, CCCP and monensin on the transport of ions across the membrane. A) The natural situation in haloarchaea shows the presence of the ions and protons either side of the membrane, with a high concentration of K^+ in the cytoplasm and a high concentration of Na^+ in the extracellular environment B) Shows the movement of H^+ caused by the presence of CCCP. C) Shows the movement of Na^+ and H^+ ions caused by the presence of monensin. This process is concentration specific and since there is much higher excess of Na^+ ions outside the cell, they are pumped into the cytoplasm to balance the levels. Since monensin is an antiporter, this causes the H^+ inside the cell to transport to the outside. The red arrows point in the direction of the ion movement.

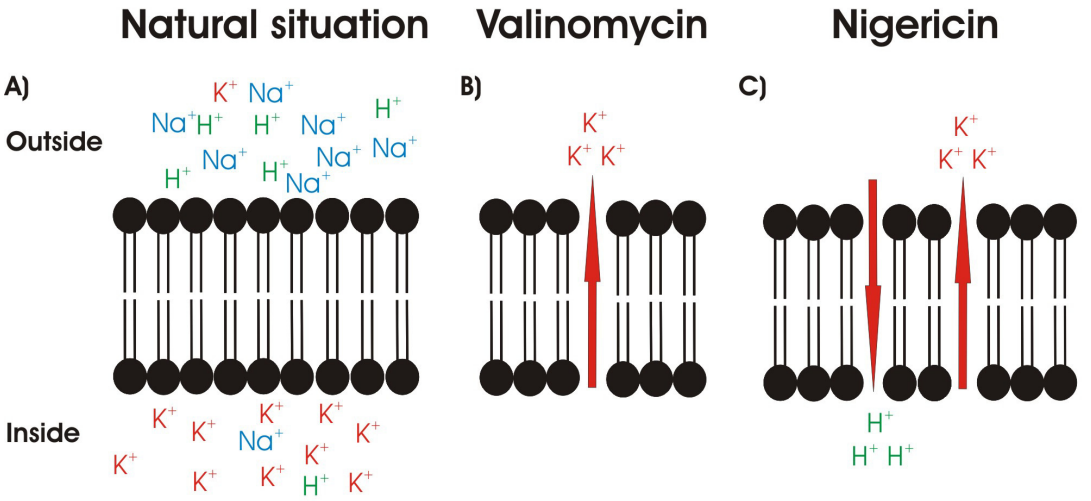

Figure 5.2 The effect of the ionophores valinomycin and nigericin on the transport of ions across the membrane of haloarchaea. A) The natural situation in haloarchaea with the presence of the ions and protons either side of the membrane B) Shows the movement of K^+ caused by the presence of valinomycin. C) Shows the movement of K^+ and H^+ ions caused by the presence of nigericin. The red arrows point in the direction of the ion movement. H^+ are hydrogen ions, K^+ are potassium ions and Na^+ are sodium ions.

Dr A. Bolhuis used the ionophores CCCP and monensin in a pulse chase experiment to show that the presence of monensin blocks translocation of AmyH while CCCP does not (Figure 5.3). This suggests that the Tat pathway is indeed driven by the SMF [152]. It was important to confirm those findings since lethal concentrations of CCCP and monensin were used, which could affect other cellular processes and thus lead to secondary effects. By determining the "minimal inhibitory concentrations" (MIC) of the ionophores on *H. hispanica* and *H. volcanii* it was possible to grow these organisms at sub-lethal concentrations. The effects of these ionophores were analyzed on the translocation of the Tat substrates, SptA and AmyH.

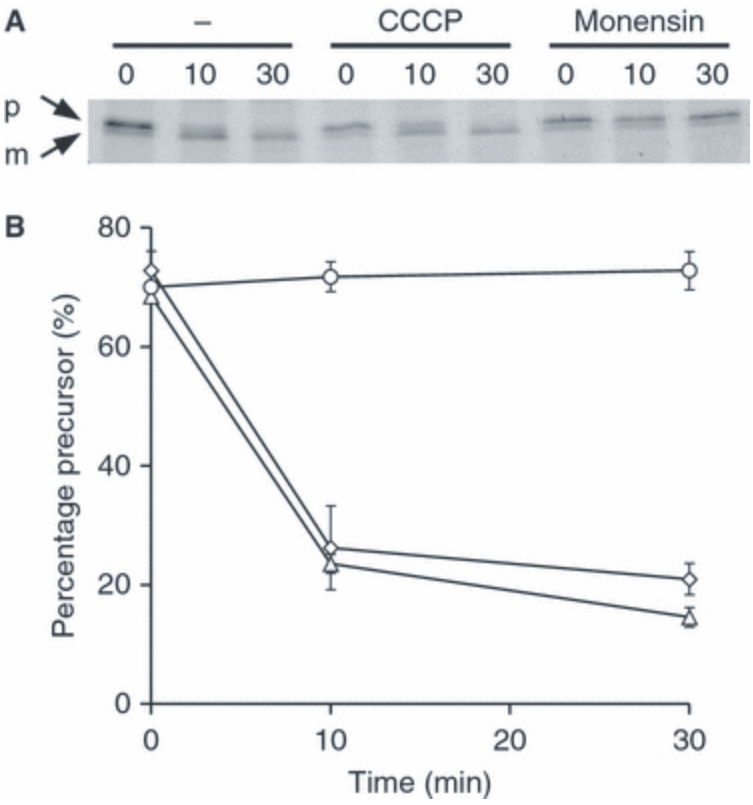

Figure 5.3 Effect of the ionophores CCCP and monensin on the translocation of AmyH. A) Pulse–chase reactions were performed on *H. hispanica* B3 in the absence or presence of CCCP or monensin. Samples were taken after 0, 10 and 30 min of chase as indicated. p, precursor; m, mature AmyH. B) The kinetics of processing were plotted as the percentage of AmyH still in the precursor form at the time of sampling. The error bars shown were calculated from two independent pulse–chase experiments. Triangles, no addition; diamonds, with CCCP; circles, with monensin. This picture was taken from Kwan *et al*., (2008) [152].

5.1.2 The effect of CCCP and monensin on the translocation of SptA via the Tat pathway

In testing the efficiency of antimicrobials, it is common practice to test them on $1x10^5$ bacterial cells per ml. There is no literature on the relationship between optical density and colony forming units (cfu) for *H. volcanii* so a viable count was performed. Results revealed that at an OD_{600} of 1 there are approximately $1.8x10^8$ cfu/ml. This allowed $1x10^5$ *H. volcanii* H26 cells per ml to be inoculated into media containing varying amounts of either CCCP or monensin. Cultures were incubated at 45°C for 3 days and results are shown in Table 5.1.

	MIC (µM)
CCCP	0.6
Monensin	2.0

Table 5.1 The MIC of monensin and CCCP for *H. volcanii* H26. Between 0.1 and 0.9µM CCCP and 0.5 to 7µM monensin were added into 4 ml of media inoculated with $1x10^5$ *H. volcanii* cells per ml. Cultures were checked for growth after 3 days and the lowest concentration where there was no growth (MIC) is shown in the table.

Next, $1x10^5$ cells per ml of *H. volcanii* (pSPTA1) were inoculated into media containing the maximum non-inhibitory concentrations of the ionophores e.g. 0.5µM CCCP or 1.5µM monensin. Cultures were incubated at 45°C for 5 days and then cells were removed by centrifugation. Finally the supernatant was concentrated 10-fold so that protease activity could be determined using azocasein. Results are shown in Table 5.2 which revealed that in the presence of monensin the activity of SptA in the supernatant was comparable to that of the wildtype. CCCP on the other hand reduced the translocation of SptA to almost non-detectable amounts.

	SptA (%)
Monensin	86.21%
CCCP	12.06%
Positive control	100.00 %

Table 5.2 The effect of monensin and CCCP on the translocation of SptA. Activity was calculated using the azocasein assay on the supernatant fraction of *H. volcanii* H26 (pSPTA1) and the background protease activity was subtracted. Calculations are shown in Appendix C and are represented in this table as a percentage with the positive control being 100% (*H. volcanii* H26 (pSPTAHIS) in the absence of ionophores). Monensin or CCCP were added at 1.5µM and 0.5µM concentrations respectively. This experiment was only performed once therefore further repeats are needed.

5.1.3 The effect of CCCP and monensin on the translocation of AmyH via the Tat pathway

A viable count of *H. hispanica* revealed that at an OD_{600} of 1 there are approximately $6x10^9$ cfu/ml. This allowed $1x10^5$ *H. hispanica* B3 cells per ml to be inoculated into HV-YPC media containing varying concentrations of monensin and CCCP. Cultures were grown for 3 days at 45°C and results are shown in Table 5.3. It was established that the highest non-inhibitory concentration of the ionophores was 1.5 µM monensin or 0.5 µM CCCP.

	MIC (µM)
CCCP	0.6
Monensin	2.0

Table 5.3 The MIC of monensin and CCCP for *H. hispanica* B3. Between 0.5 and 5 µM monensin and 0.1 to 1µM CCCP were added to 4 ml of media inoculated with $1x10^5$ *H. hispanica* B3 cells per ml. Cultures were checked for growth after 3 days and the lowest concentration of ionophore where growth was not seen (MIC) is shown in the table.

Therefore, 0.5 µM CCCP and 1.5 µM Monensin were added to HV-YPC media containing $1x10^5$ *H. hispanica* cells per ml. After 5 days incubation at 45°C the cells were harvested by centrifugation. The cells were lysed by the addition of water while the supernatant was concentrated ten-fold. Both fractions were analysed by SDS PAGE and Western blotting. Results are shown in Figure 5.4 which shows that CCCP has very little influence on the translocation of AmyH which can still be detected in both the cytoplasmic and the supernatant fractions. However, the presence of monensin seems to completely stop the translocation of the amylase.

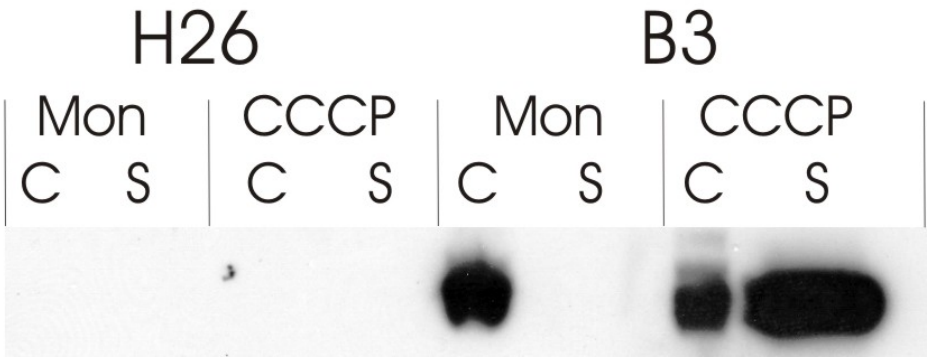

Figure 5.4 The effects of monensin and CCCP on the translocation of AmyH in *H.hispanica*. $1x10^5$ cells per ml were inoculated in HV-YPC media and grown for 5 days in the presence of 1.5 µM monensin (Mon) or 0.5 µM CCCP. The cytoplasmic and supernatant fractions were then separated and a Western blot was performed. The negative control, H26 is *H. volcanii* H26. B3 is the AmyH over expressing strain, *H. hispanica* B3. Lane C represents the cytoplasmic fraction and S represents the supernatant fraction.

5.1.4 The effect of nigericin, nonactin and valinomycin on the translocation of AmyH via the Tat pathway

It was also important to test the impact of the ionophores nigericin, nonactin and valinomycin on the translocation of AmyH. To determine the MIC of these ionophores, $1x10^5$ *H. hispanica* B3 cells per ml were inoculated into HV-YPC media containing varying concentrations of the ionophores. Nigericin was much more toxic to the cells so a much lower concentration was used. Cultures were allowed to grow for 3 days at 45°C and results are shown in Table 5.4.

	MIC (μM)
Nigericin	0.0313
Nonactin	20
Valinomycin	40

Table 5.4 The MIC of nigericin, nonactin and valinomycin on *H. hispanica* B3.
Between 0.008 and 0.0156 μM nigericin and 10 to 50 μM of nonactin and valinomycin were added to media inoculated with $1x10^5$ *H. hispanica* B3 cells per ml. Cultures were checked for growth after 3 days and the lowest concentration of ionophore where there was no growth (MIC) is shown in the table.

$1x10^5$ *H. hispanica* cells per ml were then inoculated into HV-YPC media containing the maximum non inhibitory concentrations of nigericin, nonactin and valinomycin (0.0156 μM, 15 μM and 35 μM respectively). Cultures were allowed to grow for 5 days and then the cytoplasmic and supernatant fractions were separated by centrifugation. The supernatant was concentrated ten fold and the cells were lysed by the addition of water. Both fractions were run on a SDS PAGE gel and a Western blot was performed. The results are shown in Figure 5.5 which revealed that AmyH is present in the cytoplasm and the supernatant of all four *H. hispanica* B3 cultures therefore nonactin, valinomycin and nigericin have no effect on the translocation of AmyH.

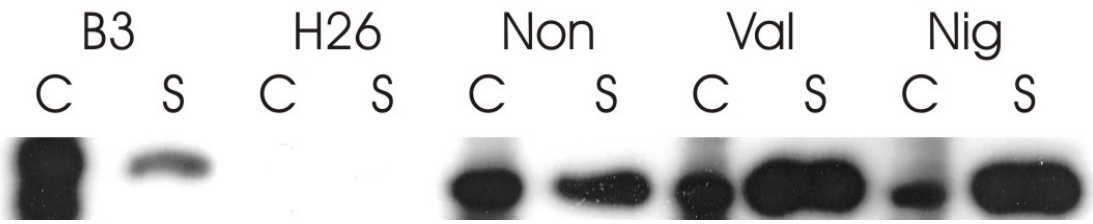

Figure 5.5 The effects of nonactin, valinomycin and nigericin on the translocation of AmyH in *H. hispanica*. *H. hispanica* B3 cultures were grown for 3 days in the presence of 0.0156 µM nigericin, 15 µM nonactin or 35 µM valinomycin (Val). The cellular (C) and supernatant (S) fractions were then resolved by SDS PAGE and a Western blot was performed. The negative control is *H. volcanii* H26 and the positive control, B3 is *H. hispanica* B3 in the absence of any ionophores. Due to the predicted high level of secretion of *H. hispanica* B3, 4 fold less was loaded onto the gel compared to the supernatant of *H. hispanica* B3 in the presence of the ionophores.

5.2 The importance of the sodium motive force for the cellular function of haloalkaliphiles

5.2.1 Introduction

Haloalkaliphiles are archaea that are akin to halophiles in their tolerance to high salt concentrations but inhabit highly alkaline environments such as soda lakes and carbonate rich soils [187]. The environmental pH ranges from pH 9-11 but analysis of cytoplasmic proteins revealed that the intra-cellular pH is closer to neutral [188]. One of the techniques that have been adopted to maintain this neutral pH is the use of Na^+/H^+ anti-porters. Sodium ions are pumped into the environment while H^+ ions are pumped back into the cell thus lowering the internal pH (Figure 5.6) [189]. Due to the highly alkaline environment this creates a reversed ΔpH compared to *E. coli*, which is compensated by the need for a large ΔΨ (See equation in introduction). Since the ΔΨ of the PMF is mainly produced by the movement of H^+ to the outside the cell it seems unlikely that the PMF drives many of the cellular processes in haloalkaliphiles [95]. It is therefore plausible that the readily available Na^+ ions contribute to the ΔΨ in the form of the SMF.

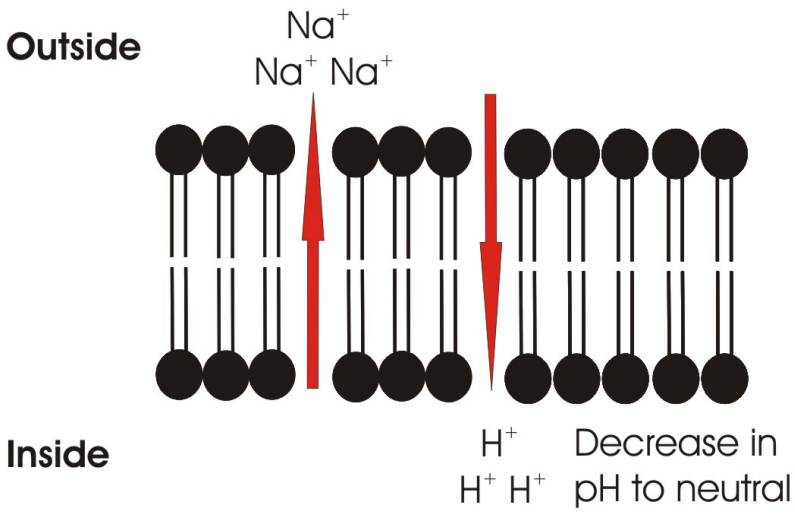

Figure 5.6 An antiporter system used by alkaliphiles to maintain a neutral intra cellular pH. The red arrows point in the direction of ion movement. H^+ are hydrogen ions and Na^+ are sodium ions.

It would be interesting to compare the MIC values of CCCP and monensin on *H. volcanii*, *H. hispanica* and the haloalkaliphiles *N. pharaonis* and *N. occultus* (See below). If the sodium gradient drives many of the haloalkaliphiles cellular features then monensin should be far more toxic to those strains as compared to the halophiles. Furthermore, visualisation of the secretome (the secreted proteome) by SDS-PAGE should show the effect of the ionophores on the organisms protein transport mechanisms.

N. pharaonis was originally isolated from Lake Gabara in Egypt and was also found in Lake Magadi in Kenya [190, 191]. It is an aerobic organism that grows optimally at 40ºC in 3.5 M NaCl, pH 8.5 although environmental pH can reach as high as pH 11 [192]. *N. occultus* was also isolated from Lake Magadi by Tindall *et al*., (1984) [190] but it has a slightly higher pH optimum then *N. pharaonis* of pH 9.5.

5.2.2 The Minimal inhibitory concentration of CCCP and monensin for a variety of haloarchaea and haloalkaliphiles.

Viable counts were performed on *N. pharaonis* and *N. occultus* which revealed that at an OD_{600} of 1 there are approximately $4.9x10^9$ and $3.8x10^6$ Cfu/ml respectively. Surprisingly, *N. occultus* had a much lower viable count which could be due to *N. pharaonis* cellular pigments producing a higher OD reading. Next, $1x10^5$ haloalkaliphile cells per ml were inoculated into media containing varying concentrations of CCCP or monensin and incubated for 3 days at 45°C. Cultures were checked for growth and results are shown in Table 5.5 (Results for *H. volcanii* and *H. hispanica* are also included).

Organism	MIC (μM) CCCP	MIC (μM) Monensin
H.volcanii H26	0.6	2
H. hispanica	0.6	2
N. pharaonis	7	0.0032
N. occultus	18	0.0015

Table 5.5 The MIC of CCCP and monensin on *H. volcanii* H26, *H. hispanica*, *N. pharaonis* and *N. occultus*. Varying concentrations of monensin and CCCP (μM) were added to media containing $1x10^5$ *H. volcanii* H26, *H. hispanica*, *N. pharaonis* or *N. occultus* cells. Cultures were checked for growth after 3 days and the lowest concentration of ionophore that growth could not be seen (MIC) is shown in the table.

The MIC results clearly showed that the haloalkaliphiles are much more resistant to the effects of CCCP compared to the halophiles, while monensin is far more toxic to their cells.

5.2.3 The effect of CCCP and monensin on the secretion of proteins from a variety of halophiles and haloalkaliphiles.

$1x10^5$ cells of *H. hispanica* B3, *H. volcanii* H26, *N. pharaonis* and *N. occultus* were inoculated into media containing the highest non inhibitory concentration of CCCP or monensin. Cultures were allowed to grow for 3 days at 45ºC and then cells were removed by centrifugation. TCA precipitation was performed on the supernatant and the purified proteins were run on a SDS PAGE gel. Proteins were visualised using the silver stain method and the results from *H. hispanica* B3 and *H. volcanii* H26 are shown in Figure 5.7.

As expected, the translocation of the Tat substrate AmyH in *H. hispanica* B3 is clearly inhibited when monensin blocks the Tat pathway (Figure 5.7, red arrow). There are also at least 9 other proteins that can no longer translocate in the presence of monensin compared to the positive control and CCCP (Figure 5.7, black arrows). Interestingly, one protein is observed only when *H. hispanica* B3 is incubated in the presence monensin (Figure 5.7, green arrow). There appears to be little difference in the secreted proteins of *H. volcanii* H26 when incubated in the presence of the ionophores. The same was also seen for *N. pharaonis* and *N. occultus* (data not shown).

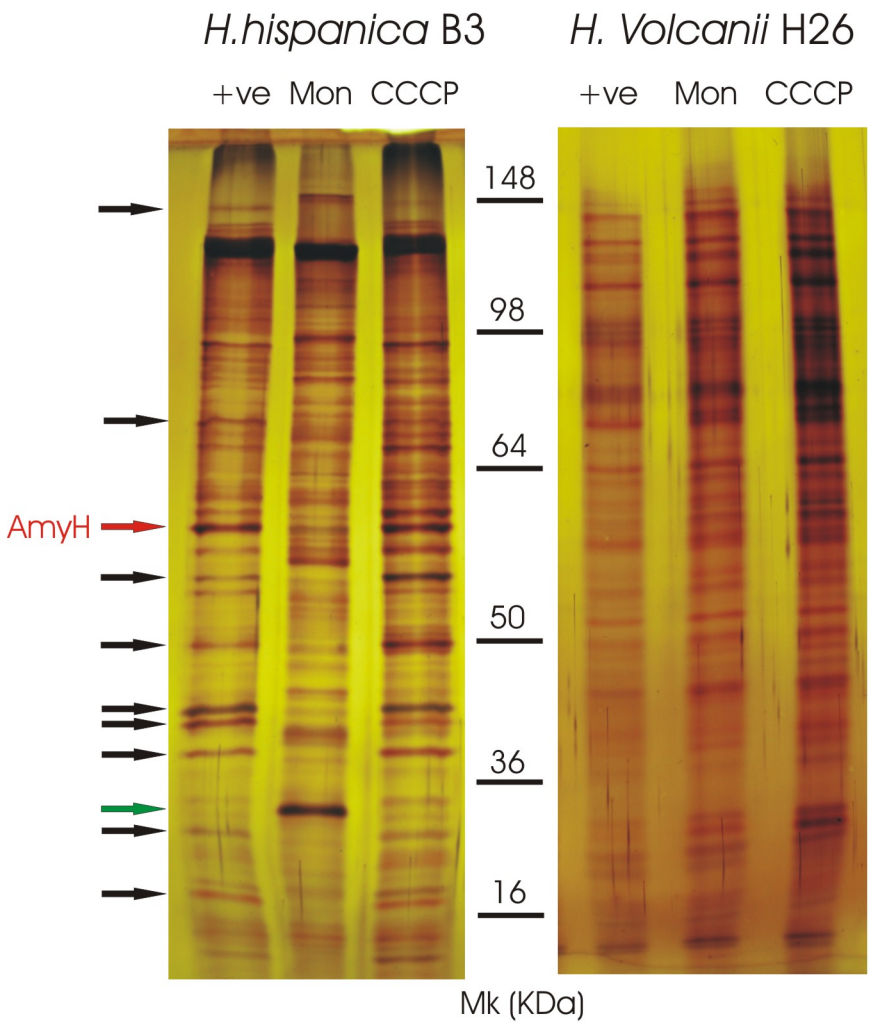

Figure 5.7 The effect of monensin and CCCP on the secretome of *H.hispanica* B3 and *H. volcanii* H26. 1x10^5 *H. hispanica* B3 and *H. volcanii* H26 cells were inoculated into HV-YPC media containing the highest non inhibitory concentration of monensin (Mon) or CCCP. Cultures were grown for 3 days at 45°C then TCA precipitation was performed on the supernatant. Samples were run on a SDS PAGE gel and a silver stain was performed. No ionophores were added to the positive control. The black arrows show proteins that appear in the positive control and the CCCP sample but not the monensin sample. The green arrow shows a protein that only appears in the monensin sample. The red arrow shows AmyH. Mk is the molecular weight marker.

5.3 Discussion

As mentioned previously, many archaea can be found in a variety of extreme environments which has caused their bioenergetics to vary greatly between species. Halophiles keep the cytoplasmic concentration of sodium ions low by using the PMF to drive H^+/Na^+ antiporters which in turn generates a high SMF [105]. Therefore, it has been suggested that this SMF could drive several of the organisms cellular features. Section 6.1 investigated whether the SMF is required for the functioning of the haloarchaeal Tat pathway. Dr A. Bolhuis has already shown using pulse chase that translocation of the Tat substrate AmyH is only affected by monensin and not CCCP suggesting that the SMF is indeed required for translocation [152].

It was important to confirm Dr. A. Bolhuis' work as he used lethal quantities of the ionophores which could lead to secondary effects. It was found that non-inhibitory concentrations of monensin did indeed block the translocation of AmyH but CCCP, valinomycin and nigericin (all disrupt components of the PMF) had no effect. This confirms that the SMF does indeed play an important role in *H. hispanica*'s Tat pathway at least for the translocation of AmyH. These results were corroborated by Dr. A. Bolhuis who used IMVs for the *in vitro* translocation of radiolabelled AmyH. He was able to show that AmyH was only translocated efficiently into the IMVs if there was sodium gradient present (produced by the high concentration of KCl outside and equimolar concentration of NaCl inside the IMV); if, however, the KCl was replaced with NaCl then translocation no longer occurred [152]. There have been other cases where cellular processes have been shown to be dependent on the SMF. For example the archaeon *Methanosarcina barkeri* requires the SMF for the oxidation of methanol [193], while both the haloarchaeon *H. salinarum* and the thermophilic bacterium *Bacillus* sp. Ta2.A1 require the SMF for the uptake of glutamate [194, 195]. To my knowledge the only findings for the involvement of the SMF in protein transport is for several of the Sec substrates in *Vibrio cholerae* and *Vibrio alginolyticus* [196, 197]. Our findings are the first evidence that the SMF is required for the secretion of a Tat-dependent substrate in *H. hispanica*.

Strikingly, the translocation of SptA when expressed in *H. volcanii* was only affected in the presence of CCCP. This suggests that it is the PMF and not the SMF that is involved in the Tat translocation of the serine protease which appears to contradict our findings with AmyH. It is a possibility that the sodium gradient is only utilised by specific proteins such as AmyH, or that it is *H. hispanica* but not *H. volcanii* that takes advantage of the available SMF for Tat translocation. It is important to note that activity measured with the azocasein assay were minimal (Appendix C) therefore it is important to perform further experiments to determine if the differences seen are significant. Unfortunately, due to time constraints it was not possible to perform a detailed study into the driving force utilised by *H. volcanii* for the translocation of the heterologous protein AmyH or Spta, it is likely that by doing so it will provide a clearer picture of the situation.

It is generally accepted that alkaliphilic bacteria generate an inwardly directed SMF which is utilised for many of their cellular processes [198]. In fact, research by Peddie *et al*., (1999) [194] have shown that the alkaliphile *Bacillus* TA2.A1 was unable to grow in as little as 0.1 μM monensin but had a high tolerance to CCCP (100 μM). This indicates the importance of the sodium gradient in alkaliphiles for the general functioning of cellular features. Similar results were found with *N. pharaonis* and *N. occultus* where monensin was incredibly toxic to the haloalkaliphiles when compared to the halophiles, *H. volcanii* and *H. hispanica*. They were also much more resistant to the presence of CCCP which gives an indication into the vital role that the SMF plays in driving haloalkaliphilic cellular processes.

As mentioned above, the translocation of AmyH *via* the Tat pathway of *H. hispanica* appears to be driven by the SMF which also plays an important role in processes within haloalkaliphiles. Therefore, the results from the analysis of the organisms secretome was surprising as you would expect there to be a greater effect produced by the dissipation of the sodium gradient by monensin. For example, more then 60 % of halophilic proteins are predicted to be Tat substrates, yet there were only 9 *H. hispanica* proteins whose translocation could visually be seen to be inhibited by the ionophore. The experiment worked correctly as the translocation of AmyH is clearly inhibited when the Tat pathway is blocked. There were no obvious differences in the secretion of proteins by the haloalkaliphiles and *H. volcanii* in the presence of these

ionophores. Therefore, it is likely that higher concentrations of the ionophores are required to visualise their effects. However, higher concentrations will be above the MIC value and thus will inhibit growth or kill the cells. This type of experiment could only work with a pulse chase as this technique only analyses the synthesis and secretion of proteins during a very short time period. Unfortunately, the amounts of protein secreted were apparently too low for visualisation using pulse chase. It might be possible to include a protein concentration step (e.g. through TCA precipitation), but due to time constraints this experiment has not been done.

Chapter 6: The stability and folding of the Tat dependent protein AmyH

6.1 Characterising AmyH from *H. hispanica*

6.1.1 Introduction

Haloarchaea are the only organisms that transport the majority of their proteins *via* the Tat pathway [21, 118]. Most of these proteins do not bind to co-factors so there must be other reasons for folding before translocation. The first reason is most likely to prevent aggregation caused by the high levels of KCl in the cytoplasm [114]. By folding rapidly it means that the hydrophobic residues are quickly buried while the acidic residue are exposed causing electrostatic repulsion (See introduction) [120, 122]. It is also possible that the high levels of salt have meant that chaperones may be required for the correct folding of halophilic proteins [123]. If these chaperones are cytoplasmic then proteins would have to fold before translocation occurs.

There have been relatively few starch degrading amylases discovered in haloarchaea and only five of these have been characterised in any detail [150]. These amylases originated from *Natronococcus amylolyticus* sp. strain Ah-36 [199], *H. salinarum* [200], *Haloferax mediterranei* [201] and *Haloarcula sp.* S-1 [202]. Only the *N. amylolyticus* genome has been sequenced and Rose *et al*., (2002) have shown using mutagenesis of the RR-motif that the amylase is a Tat substrate [118, 203]. Chapter 4 also confirmed that the amylase AmyH from *H. hispanica* is a Tat substrate. These amylases do not bind to co-factors therefore by characterising AmyH it may provide some information into why so many halophilic protein fold before translocation.

After purification of AmyH, there are two main areas of investigation that will be focused on. The first is to investigate the stability of the amylase while the second is to determine what conditions are required for the efficient folding of AmyH. It is important to perform these experiments since the rate and timing of folding and also the stability of proteins is intimately linked to their secretion

6.1.2 Purification of AmyH

To purify AmyH, 2 litres of Hv-YPC was inoculated with *H. hispanica* B3 and grown for six days. This particular *H. hispanica* strain has significantly higher levels of AmyH secretion compared to that of the wildtype [150]. Cells were removed by centrifugation and the supernatant was centrifuged for a second time to remove any remaining cells. The sample was then concentrated 20-fold using the Vivaflow 200 tangential filter unit.

The concentrated supernatant was passed through a β-cyclodextrin-sepharose column allowing AmyH to be purified. 1mM dithiothreitol (DTT) was added to all purification buffers to maintain a reducing environment. AmyH was eluted in 1.5 ml fractions which were then run on a SDS PAGE gel. The gel was stained with Coomassie and the results in Figure 6.1 showed that AmyH was present in fractions 5 to 21. These samples were pooled together and the β-cyclodextrin was removed by dialysis.

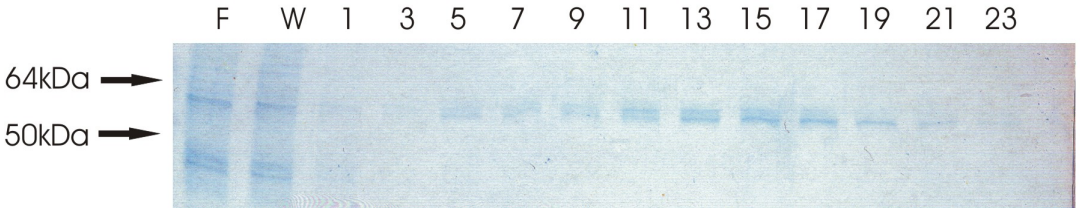

Figure 6.1 Purification of AmyH. AmyH was purified on a β-cyclodextrin-sepharose column and then the resulting fractions were run on a SDS PAGE gel and resolved by Coomassie stain. Lane F shows the initial flow through from the column and Lane W shows the flow through from the column when washed with Wash Buffer. Lanes 1-23 shows the separate 1.5 ml elution fractions containing the purified AmyH.

6.1.3 The temperature stability of AmyH

Hutcheon *et al*., (2005) [150] previously determined that purified AmyH has an optimum activity at 50°C in 4-5M NaCl buffer pH 6.5. Therefore the stability of AmyH at different temperatures was measured to determine if there is any difference when incubated in the presence of KCl or NaCl. 10 µl of purified amylase was added to Buffer (50 mM Tris pH 7.5, 5 mM $CaCl_2$ and 10 mM DTT) containing either 4 M NaCl or 4 M KCl. A PCR cycling machine was used to incubate these samples at temperatures between 55ºC and 75°C. After 20 minutes the Phadebas amylase assay was performed and results are shown in Figure 6.2.

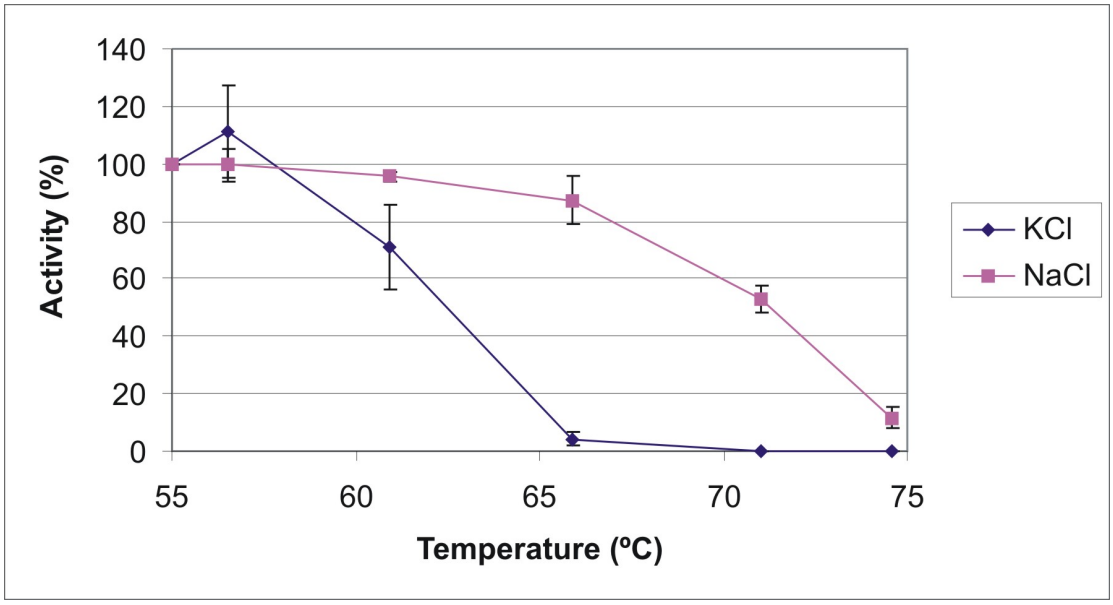

Figure 6.2 The thermostability of AmyH in NaCl and KCl. The x-axis shows the different temperatures that AmyH was incubated at for 20 minutes before the Phadebas amylase was performed. The y-axis shows the specific activity of these AmyH samples, the activity at 55ºC was taken as 100% activity. The pink line represents AmyH incubated in 4 M NaCl while the blue line represents AmyH incubated in 4 M KCl. The data for this graph are shown in Appendix D.

Results show that the activity of AmyH incubated in KCl rapidly decreased as temperature rose above 56ºC. In fact at 65°C no amylase activity could be detected. Interestingly, the decrease in activity of AmyH incubated in NaCl was much more gradual. The protein maintained high activity at 65ºC; even at 71ºC there was still 50% activity, while at 75ºC the activity dropped to below 20%.

6.1.4 Stability of AmyH at 50°C and 70°C

The stability of purified AmyH at 50°C and 70°C was analysed in more detail. First, 10 µl of purified AmyH was incubated in Buffer (50 mM Tris pH 7.5, 5 mM $CaCl_2$ and 10 mM DTT) containing either 4 M NaCl or 4 M KCl. Samples were incubated at 50ºC and the Phadebas amylase assay was performed every hour for 5 hours. The results are shown in Figure 6.3 which revealed that there is no significant decrease in activity over the five hour period.

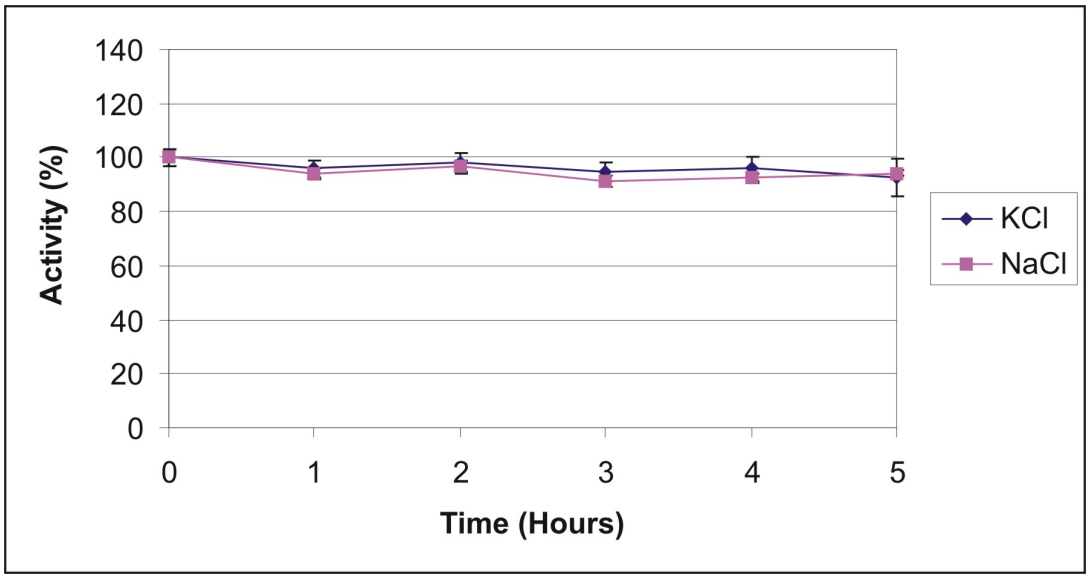

Figure 6.3 The stability of AmyH at 50ºC in either NaCl or KCl. The x-axis is the amount of time that AmyH was incubated at 50ºC before the Phadebas amylase assay is performed. The y-axis shows the specific activity of AmyH which is calculated using activity at time 0 as 100%. The pink line represents AmyH incubated in 4 M NaCl while the blue line is AmyH incubated 4 M KCl. The data for this graph is shown in Appendix D.

To investigate the stability of AmyH at 70ºC, 10µl of the purified AmyH was incubated in Buffer (50 mM Tris pH 7.5, 5 mM $CaCl_2$ and 10 mM DTT) containing either 4 M NaCl or 4 M KCl. Samples were incubated at 70ºC and activity was measured every 10 minutes for 1 hour using the Phadebas amylase assay. Results are shown in Figure 6.4.

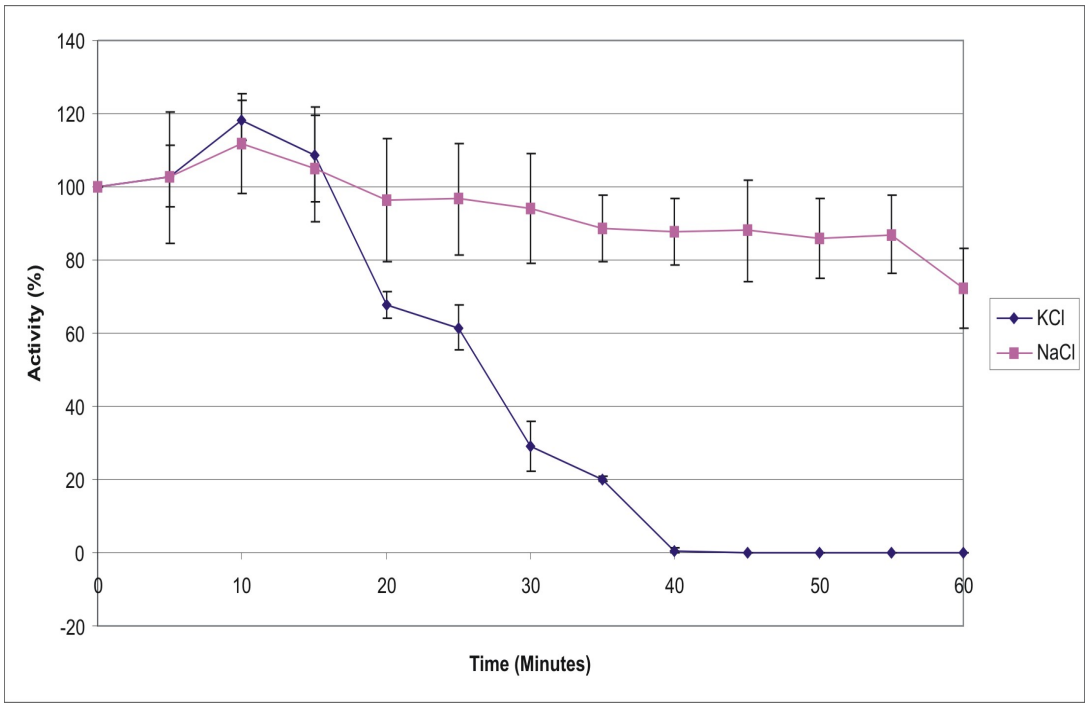

Figure 6.4 The stability of AmyH at 70ºC in either NaCl or KCl. The x-axis shows the time in minutes that AmyH was incubated at 45°C before the Phadebas amylase assay was performed. The y-axis shows the specific activity of AmyH which was calculated using activity at time 0 as 100%. The pink line represents AmyH incubated in NaCl while the blue line represents AmyH incubated in KCl. The data for this graph is shown in Appendix D.

Similar to the results in Figure6.2 the activity of AmyH incubated at 70ºC in KCl rapidly decreased after 15 minutes and is no longer detectable after 40 minutes. In contrast, in NaCl the decrease in activity is much more gradual with there still being 75% activity after 60 minutes.

6.1.5 Time dependent refolding of AmyH

It is important to determine optimal conditions for refolding of AmyH; therefore, purified AmyH was unfolded by dialysis in 50 mM Bis Tris buffer pH 6.5 containing 6M urea. It must be noted that during these refolding trials, refolding efficiency of AmyH was not determined. However, Dr Bolhuis has previously compared this and showed that under optimal conditions the refolding efficiency is at least 80-85% (Unpublished data). Even under suboptimal conditions, e.g. in the absence of salt, refolding efficiency was still 60% (Kwan *et al*., 2008). Thus, results from these refolding assays can be assumed to be significant.

To test the amount of time it takes for this AmyH to refold into an active conformation, a folding assay was performed using 10 µl of unfolded AmyH in 200 µl of buffer (50 mM Tris pH 7.5, 5 mM $CaCl_2$, 10 mM DTT and 3.5 M KCl). This dilutes the urea thus allowing refolding to occur. Samples were incubated at 45°C and an amylase assay was performed after set time points. The results are shown in Figure 6.5.

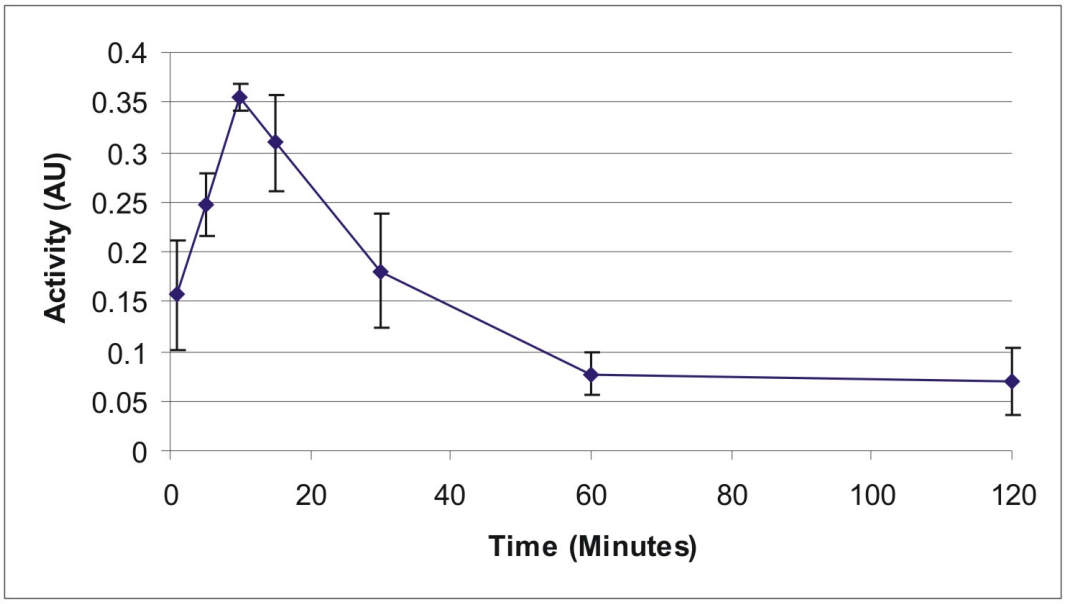

Figure 6.5 Time-dependent refolding of AmyH. The x-axis is the period of time in minutes that AmyH was allowed to refold at 45ºC. The Phadebas amylase test was then performed and the resulting activity is shown on the y-axis. Activity was measured in arbitrary units (AU). The data for this graph is shown in Appendix D.

The assay revealed that AmyH folds rapidly after being added to the Buffer with the highest activity detected after 10 minutes. Surprisingly, activity of AmyH rapidly decreased again if the refolding reaction was performed longer then 15 minutes.

6.1.6 The requirement of $CaCl_2$ for folding.

Hutcheon *et al*., (2005) [150] have shown that $CaCl_2$ is required for AmyH activity so to determine if it is required for refolding, 10 µl of unfolded AmyH was added to 200 µl of buffer (50 mM Tris pH 7.5, 10 mM DTT and 3.5 M KCl). The samples were incubated at 45°C with or without the presence of 5 mM $CaCl_2$. After 10 minutes the amylase assay was performed and the results are shown in Table 6.1.

	Activity
+ $CaCl_2$	100%
- $CaCl_2$	98.19%

Table 6.1 Activity of AmyH after refolding in the presence or absence of 5 mM $CaCl_2$. The data for this table are shown in Appendix D.

The results indicate that the presence of $CaCl_2$ has no effect on the folding of AmyH. Note that although it is not required for folding, it is added to all folding and stability buffers as it is required for AmyH activity.

6.1.7 The need of a reducing environment for AmyH refolding

As discussed in Chapter 7.2 below, it is likely that the *H. volcanii* cytoplasm is a reducing environment so it is important to investigate if reducing conditions are required for the folding of AmyH [204]. 10 µl of unfolded amylase was added to 200 µl of Buffer (50 mM Tris pH 7.5, 5 mM $CaCl_2$ and 3.5 M KCl) containing varying amounts of the reducing agent DTT. The samples were incubated at 45°C for 10 minutes to allow refolding to occur. The Phadebas amylase assay was then performed and results are shown in Figure 6.6.

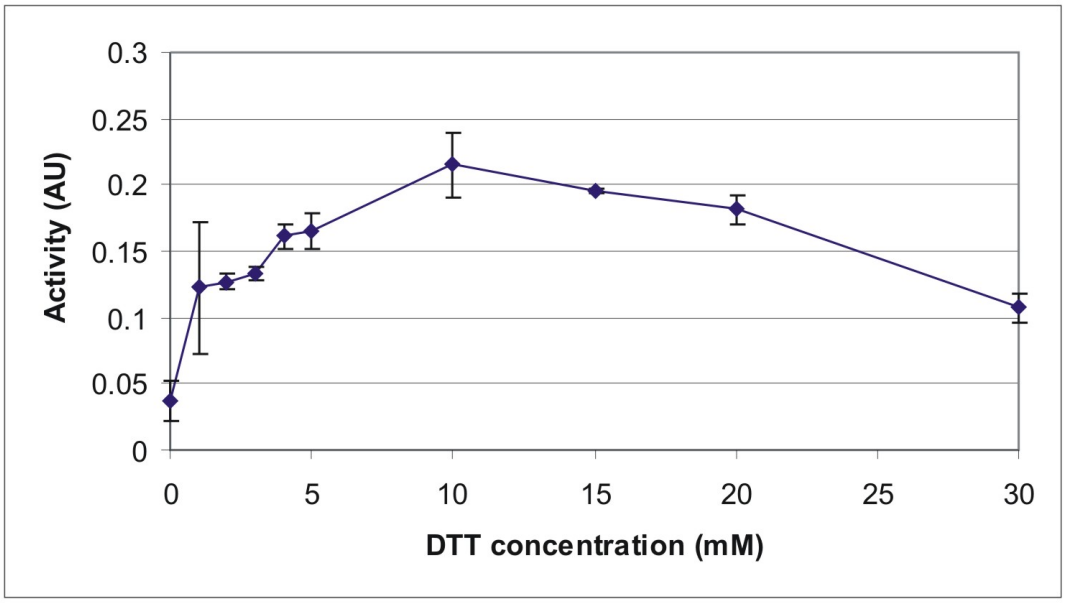

Figure 6.6 Dependence of AmyH refolding on DTT concentration. The x-axis shows the concentration of DTT that was added to the unfolded AmyH. After 10 minutes incubation at 45ºC the Phadebas amylase assay was performed. The y-axis shows the resulting activity which was measured in arbitrary units (AU). The data for this graph is shown in Appendix D.

AmyH activity was barely detectible in the absence of DTT but increased considerably when as little as 1 mM was added. Therefore, this indicates that a reducing environment is important for folding. Activity continues to increase with DTT concentration until the optimum is reached around 10 mM. If more then 20 mM DTT is added then activity of AmyH begins to decrease again.

6.1.8 The pH optimum for re-folding of AmyH

To determine the most efficient pH for AmyH folding, 10 µl samples of unfolded protein were incubated in 200 µl Buffer (5 mM $CaCl_2$, 10 mM DTT and 3.5 M KCl) containing 50 mM acetate-NaOH (pH 3-5.5), 50 mM BisTris-HCl (pH 6-7) or 50 mM Tris-HCl (pH 7.5-9),. Samples were incubated for 10 minutes at 45ºC and then the Phadebas amylase assay was performed, the results are shown in Figure 6.7.

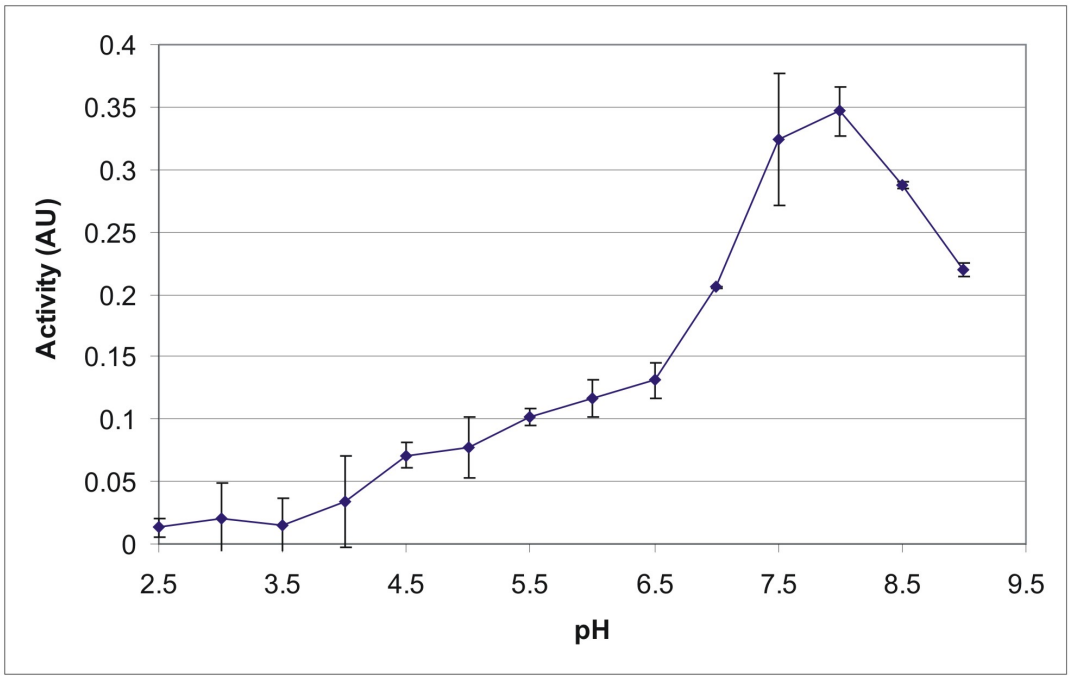

Figure 6.7 pH dependence of the refolding of AmyH. The x-axis shows the different pH that the unfolded AmyH was incubated in. After 10 minutes at 45ºC the Phadebas amylase assay was performed. The y-axis shows the resulting activity which was measured in arbitrary units (AU). The data for this graph are shown in Appendix D.

The graph indicates that there was very little refolding between pH 2.5 and 4 but that efficiency of refolding gradually increased between pH 4.5 and 6.5. Efficient folding began to occur as pH rose above 6.5 with the highest activity detected between pH 7.5 and 8. Any further increase in pH began to be detrimental to folding again.

6.1.9 The NaCl and KCl concentrations required for AmyH to re-fold.

Due to the high extra-cellular NaCl and intra-cellular KCl concentrations found in *H. volcanii*, the ability of AmyH to fold in these salts was tested. 10 µl of unfolded amylase was added to 200 µl of Buffer (50 mM Tris pH 7.5, 5 mM $CaCl_2$ and 10 mM DTT) containing varying concentrations of either NaCl or KCl (0.5M to 4.5 M in 0.5 M intervals). Reactions were incubated for 10 minutes at 45ºC and then the Phadebas amylase assay was performed. The results are shown in Figure 6.8.

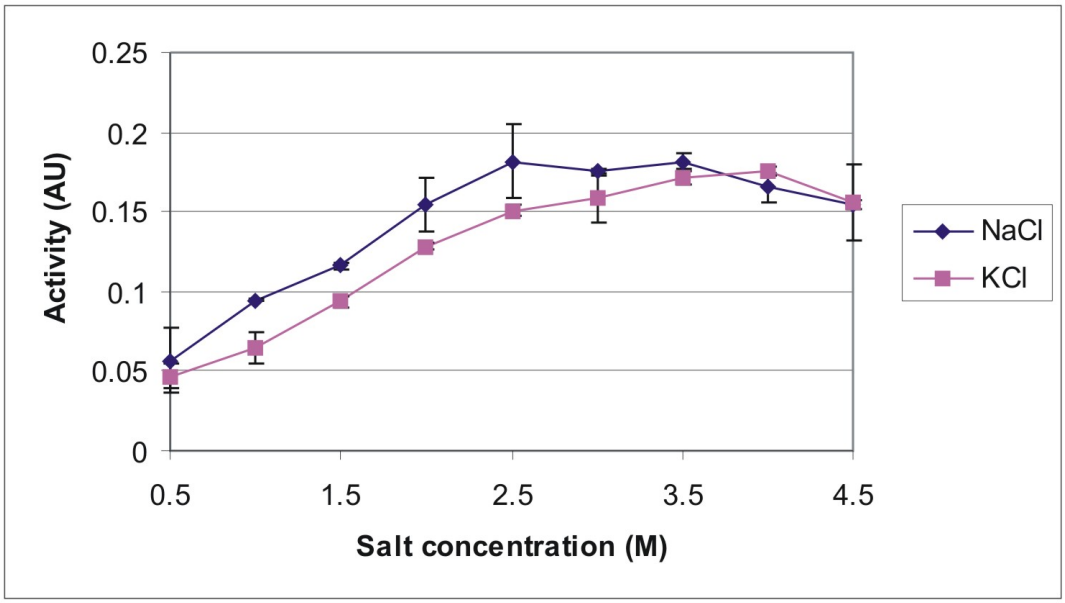

Figure 6.8 Salt dependence of the refolding of AmyH. The x-axis shows the salt concentration of NaCl and KCl that the unfolded AmyH was incubated in. After 10 minutes the Phadebas amylase assay was performed and the resulting activity was measured in arbitrary units (AU) as represented on the y-axis. The purple line represents AmyH incubated in KCl while the blue line represents AmyH in the presence of NaCl. The data for this graph are shown in Appendix D.

At lower concentrations, AmyH seems to have a slightly higher activity in NaCl but there is very little difference between the two salts. Interestingly, protein folding still occurs at salt concentrations as low as 0.5 M. Maximum activity is reached when incubated at salt concentrations of 2.5 M or higher for both salts.

6.2 Discussion

The aim of Chapter 6 was to characterise the *H. hispanica* α-amylase, AmyH. It was hoped that by doing so it would help us to gain a better understanding of the folding and stability of the protein, processes that are intricately linked to Tat-dependent protein translocation. AmyH was successfully purified from the supernatant of the overproducing strain, *H. hispanica* B3 therefore allowing for further experiments.

The amylases from *N. amylolyticus* and *H. mediterranei* have optimal activity at 55°C but are no longer active at 60ºC. Mijts and Patel (2002) [205] describes this as a common property of halophilic amylases. AmyH is slightly different because it has a lower optimum of 50°C but activity is not lost until 70°C [150]. It was expected that similar to activity, AmyH would be most stable around 50ºC and this turned out to be the case. What was not expected was that as temperature increased above 50ºC then activity of AmyH in KCl dropped much more rapidly than when in NaCl. One possible explanation is that AmyH is secreted and functions in the environment, and therefore it makes sense for the protein to be more stable in NaCl [206]. This is corroborated by the fact that there are several examples of halophilic cytoplasmic proteins that are more stable in KCl then NaCl such as succinic dehydrogenase from *H. salinarium* and a ribosomal protein from *Halobacterium cutirubrum* [207, 208]. However, there have also been found to be several cytoplasmic halophilic proteins that contradict this theory such as glutamine synthetase and glucose dehydrogenase from *H. mediterranei* [209, 210] and HmMalDH from *H. marismortui* [211]. These proteins have been found to be more stable in NaCl than KCl; however, it has been suggested by Tehei *et al*., (2001) [211] that this is because of the higher hydration and binding energies of Na^+ compared to K^+.

AmyH was unfolded in 6 M urea so that a variety of re-folding assays could be performed [150]. These assays measured the activity of AmyH which is ultimately linked to the amylases ability to refold. When the urea was diluted then this allowed unfolded AmyH to refold with the highest resulting activity being detected between 10 and 20 minutes incubation. What was surprising was that if allowed to refold for longer than 20 minutes then AmyH quickly lost its activity again. Under the same

conditions, purified AmyH that was not treated with urea was very stable - not loosing any activity over a 5-hour period. From this we can infer that it is the refolding process that is leading to its instability. One possibility is that stabilising factors of AmyH were co-purified but became inactive during the unfolding process, which might explain the instability of AmyH when allowed to refold for long periods of time. Another possibility is that there are two processes occurring simultaneously, the folding and the aggregation of unfolded/folding intermediates of AmyH. If folding is the more rapid process then this explains the initial activity but eventually, the aggregation of the unfolded/folding intermediates may in-turn lead to the aggregation of active amylase and consequently loss of activity.

H. hispanica survives in a very saline environment that contains at least 15% NaCl with optimum growth occurring at 25% NaCl (~4-5 M) [126]. To overcome the osmotic pressure the cell accumulates 4-5 M KCl in the cytoplasm using the "Salt in-strategy" [114]. The extracellular pH of *H.hispanica* varies between pH 6 and pH 8 (optimum pH 7) with the intercellular pH being predicted to be around neutral [126, 212]. It is of no surprise that Hutcheon *et al*., (2005) [150] discovered that its optimum activity of AmyH (4M NaCl or 4 M KCl at pH 6.5-7) occurred in conditions similar to the internal and external conditions of the organism. It was predicted that since AmyH folds naturally in the cytoplasm it would fold most efficiently when incubated in KCl at pH 7. This was not the case as results showed that there is little difference between NaCl and KCl with optimum activity being detected between 2.5 M and 4.5 M salt. Optimum folding occurred at a slightly higher then expected pH of pH 7.5 - 8.5.

In general, halophilic proteins become permanently inactive if NaCl concentration drops below 2 M [116]. Interestingly, it was discovered that the α-amylases from *N. amylolyticus* [199], *Haloarcula sp.* S-1 [202] *H. mediterranei* [201] and *H. salinarum* [200] can function at lower NaCl concentrations (as low as 1.7 M for *Haloarcula sp.* S-1 and 1.3 M for the other Haloarchaea). D'Souza *et al*., (1997) [213] suggested that this is because these amylases are secreted so have to be able to withstand changes in salt concentration caused by, for instance, rain falling into the salt lakes where haloarchaea are found. If so, one would expect a similar stability at low salt concentrations for other secretory proteins from haloarchaea. Indeed, the protease

SptA from *Natrinema* sp. 17 was shown to maintain 70% activity when the salt concentration dropped as low as 0.2 M [172]. Hutcheon *et al*., (2005) [150] have already shown that AmyH can function in much lower salt concentrations then the other halophilic amylases investigated with 30% activity still detected between 0 and 1 M NaCl. This was also found to be true for the folding of AmyH where ~ 29 % of the activity was still detected in 0.5 M NaCl or 0.5 M KCl. Unlike halophilic cytoplasmic proteins which are obligatorily salt dependent, it appears that AmyH and SptA are both stable at low salt concentrations. This suggests the possibility that this is perhaps a general feature of halophilic secretory proteins thus allowing them to tolerate changes in the environmental conditions.

Gupta *et al*., (2003) [214] and van der Maarel *et al*., (2002) [215] discuss that many α-amylases require the binding of Ca^{2+} ions to remain active. Hutcheon *et al*., (2005) [150] showed that this is true for AmyH with ~ 100% activity detected in the presence of 2 to 6 mM $CaCl_2$ but only 30% activity when $CaCl_2$ is absent. Although $CaCl_2$ is required for activity it appears that its presence has very little effect on the folding of the amylase.

The *E. coli* cytoplasm is a reducing environment which is maintained by proteins such as thioredoxin. Since disulphide bonds cannot form in these reducing conditions they tend to only be found within secreted proteins [216]. For this reason it is predicted that most other prokaryotes also have a reducing cytoplasm since they encode cytoplasmic thioredoxins and disulphide forming secretory proteins [217, 218]. To our knowledge there are only a few thermophilic prokaryotes that do not follow this rule since they use disulfide bonds to allow cytoplasmic proteins to remain stable at these high temperatures [219]. DTT (a reducing agent) has been added to all previous folding assays to maintain a reducing environment. It was important to determine whether these reducing conditions are an influencing factor on AmyH folding in the cytoplasm. It was found that 1 mM DTT would double and 10 mM DTT (optimum activity) would quadruple activity compared to unfolded AmyH lacking DTT. This confirms that a reducing environment is important for AmyH folding in the cytoplasm. One possible reason is because the amylase possesses two free cystines that, if fully exposed to an oxidising environment, could form intra- or inter-molecular disulphide bonds, leading to incorrect folding and loss of activity. By

folding in the reducing environment of the cytoplasm, prior to translocation, the protein will be more stable with less chance of these cysteines interacting to cause aggregation when in the environment.

To conclude, the purification of AmyH allowed a variety of folding and stability experiments to be performed. Results were mainly as expected with AmyH most efficiently folding in cytoplasmic conditions and most stable at 50°C. The two main results of note were that AmyH requires a reducing environment for efficient folding and is more stable in NaCl than in KCl.

Chapter 7: Final Discussion

7.1 Discussion

To conclude, this thesis has focused on the twin arginine translocation pathway which has intrigued many scientists in its ability to translocate fully folded proteins across the membrane. There has already been extensive research into this area in the bacterial and chloroplast systems although many questions still remain, the main one being the actual mechanism of translocation. The Tat pathway in archaea is still a relatively new area of research that could potentially unlock the answers to many of these questions. Therefore, this thesis has focused on thermophilic and halophilic archaea with the view that their extreme habitats will have produced unique adaptations to the Tat pathway.

The initial work in this thesis investigated the Tat pathway found in the thermophilic acidophile *S. solfataricus*. This organism was of particular interest because, unlike other archaea investigated (including other *Sulfolobus* species), *S. solfataricus* has two separate Tat operons, a situation that is reminiscent of Gram positive bacteria. The initial investigation involved the purification of both Tat complexes in *E. coli* in the hope that it would reveal more into their interactions. Unfortunately, although the TatA proteins could be purified the two TatC proteins appeared to be toxic to *E.coli*. Further experiments were performed on the purified TatA proteins with the two results of note being that they form homo-oligomeric complexes but, as expected, were unable to complement the *E. coli* Tat system. Ideally with more time purified TatA could also be used to raise antibodies and for crystallisation trials, both of which would potentially reveal much more information about this system. It was always intended to purify the *S. solfataricus* Tat complexes in their native organism but unfortunately this proved difficult with the problems with cloning and expression, and therefore further work is needed.

The other part of the *Sulfolobus* project looked at individually knocking out the two Tat operons found in *S. solfataricus*. It was hoped that this would reveal the importance of the two Tat complexes and also if there is protein specificity as is found in several Gram positive bacteria [71]. Although transformation was possible using the *S. solfataricus* gene knockout technique developed by Albers and Driessen

(2007) [132], knockouts could not be obtained which suggests that the Tat pathway may be an essential process in *S. solfataricus*. This would be an interesting result as *H. volcanii and H. salinarium* are the only known archaea to require a functioning Tat pathway [119]. One possible way of confirming that the Tat pathway is essential would be to transform a second plasmid encoding the Tat operons behind an inducible promoter. Then it should be possible to successfully knockout the chromosomal Tat operon. If the Tat pathway is indeed essential then depletion of the plasmid expressed Tat proteins should then inhibit cell growth.

The rest of this thesis focused mainly on the Tat pathway of haloarchaea where this system has been found to be the major route of translocation [117, 118]. Initial work determined that the twin arginines from the Tat motif of SptA and AmyH are essential for the substrates to translocate. By mutating one or both arginines to lysines would result in a block in translocation. Interestingly, it appears that the AmyH-KR (and also the AmyH (V17A) mutation) but not the SptA-KR mutant increases cell lysis/translocation in *H. volcanii* when grown at room temperature. This warrants further investigation to determine why this specific mutation makes such a difference. For example, if Amy-KR is becoming stuck in the translocase leading to cell lysis then cross-linking can be used in conjunction with co-purification to determine if the TatA or TatC is in close proximity. This experiment of course will not work if AmyH-KR is binding to the membrane and not the translocase. It is also important to exchange the AmyH signal peptide with that of the SptA (R11K) signal peptide (which does not produce a halo at room temperature) to determine if this affects the results seen when grown at room temperature.

Mutation of the other 5 amino acids of the twin arginine motif of AmyH to alanine residues revealed very similar results to that of the Tat substrate SufI from *E. coli* [174]. The only difference was that mutating the leucine to an alanine (L18A) blocked translocation of AmyH but the mutation had little effect on the translocation of SufI. The Seqlogos also suggest the importance of this leucine for other halophilic Tat substrates. Ideally, with more time a similar experiment would be performed on SptA to see if these results correspond to the AmyH findings. These plate assays have proved to be a useful tool that potentially could be used to look into further detail at the AmyH Tat signal peptide. For example, is translocation affected if the H-domain

is made more hydrophobic akin to the Sec signal peptide, this proved true when performed on the TorA signal peptide from *E. coli* (replaced the hydrophobic H-domain with a more hydrophobic region which caused the protein to be rerouted *via* the Sec pathway [32]). Also what are the effects of mutating the positively charged residues at the end of the C-domain that are thought to contribute to the Sec-avoidance system [34].

Further work that utilised AmyH involved investigating the driving force for the Tat pathway in *H. hispanica*. Using a variety of ionophores it became clear that the system requires a sodium gradient and therefore utilises the SMF to drive translocation of AmyH. This differs from the finding of the bacterial and chloroplast systems which have been shown to be driven by the PMF [58]. Interestingly, initial work into SptA suggests that it requires the PMF to drive the Tat pathway which again indicates the possibility that *H. volcanii* and the *H. hispanica* Tat systems differ in the way they function. Thus, it is important to confirm that other *H. hispanica* Tat substrates require the SMF for translocation and if it is the SMF or the PMF that drives the translocation of AmyH in *H. volcanii*. It conjunction to this, it is also important to determine if SptA does indeed require the PMF for translocation in *H. volcanii*.

Ionophores were also used to investigate the bioenergetics of the archaeal species, haloalkaliphiles. The lethality of monensin to *N. pharaonis* and *N. occultus* hinted at the possible importance of the SMF to these organisms. Analysis of the secretome was also performed to provide a clearer picture into the effects of monensin on the organisms protein translocation. Unfortunately, the toxicity of the ionophore and possibly the high amount of background proteins meant that no clear difference could be seen. A Pulse Chase experiment should be performed since it will overcome the toxicity issue and provide a clearer picture into the effect of the ionophores on the secretome. However, previous attempts have revealed the need to concentrate the secreted proteins in the supernatant thus a technique such as TCA precipitation will need to be included.

Finally, AmyH was purified allowing it to be characterised using a variety of folding and stability assays. The majority of results were as predicted with AmyH being

most stable in conditions resembling the extra-cellular environment and optimum folding occurring in conditions similar to the intra-cellular environment. The main result of note is that AmyH requires the reducing cytoplasmic environment to fold, therefore this is one of most likely several influencing factors causing AmyH to require the use of the Tat pathway. What was surprising was that refolded AmyH became inactive if it was allowed to refold for longer then 15 minutes. It is possible that this is because the mixed population of folded/unfolded amylase aggregated overtime which caused activity to be lost. To test this theory, unfolded AmyH should be allowed to refold in lower salt concentrations or in the presence of a detergent such as Triton X-100. Both of which should reduce the rate of aggregation and thus allow AmyH to maintain activity when allowed to fold for longer then 20 minutes. Finally, with more time there are still several experiments that should be performed to characterise AmyH in more detail. For example the stability of AmyH in the presence/absence of DTT and the ability of the amylase to fold in the absence of salt should be tested.

In Summary, the Tat pathway has provided an intriguing area of research into the translocation of fully folded proteins across the membrane. There have been many discoveries into the bacterial and the equivalent chloroplast systems although there is still much to be learnt about how this system functions. This thesis has presented an in-depth investigation into the relatively new area of the Tat pathway in archaeal species and has specifically looked at the process within sulfolobus and haloarchaea. The resulting work represents a substantial and original contribution to knowledge that will hopefully further our understanding into the Tat pathway.

Appendix A: Supplementary results from Chapter 3

Found below is the data relevant to the work reported in Chapter 3 which is represented by graphs within the main body of text.

Time	$pSA_1A_2C_1$ - ara	$pSA_1A_2C_1$+ ara	pSA_3C_2 - ara	pSA_3C_2+ ara
0	0.000	0.000	0.000	0.000
1	0.066	0.024	0.080	0.029
2	0.181	0.063	0.188	0.080
3	0.358	0.224	0.408	0.245
4	0.509	0.26	0.625	0.259
5	0.626	0.284	0.769	0.262
6	0.700	0.284	0.805	0.260

Table A.1 Data for the growth curve of *E. coli* JM109 transformed with $pSA_1A_2C_1$ and pSA_3C_2 (Figure 3.3). Arabinose was added after 3 hours and the absorbance was measured every hour at OD_{600}.

Time	pSC_1 - ara	pSC_1+ ara	pSC_2 - ara	pSC_2+ ara
0	0.000	0.000	0.000	0.000
1	0.024	0.026	0.008	0.014
2	0.068	0.078	0.037	0.04
3	0.122	0.123	0.105	0.111
4	0.189	0.326	0.334	0.369
5	0.211	0.629	0.39	0.654
6	0.251	0.947	0.411	0.913

Table A.2 Data for the growth curve of *E. coli* JM109 transformed with pSC_1 and pSC_2 (Figure 3.6). Arabinose was added after 3 hours and the absorbance was measured every hour at OD_{600}.

Appendix B: Supplementary results from Chapter 4

Found below are data tables relevant to the work reported in Chapter 4.

Protein	Signal peptide length	Length of H domain	Hydrophobicity of H domain	Average Hydrophobicity
Cbin	27	21	41	1.95
Csg	34	22	37.8	1.72
Edp	33	19	38	2.00
HemV1	24	16	32.9	2.06
Vng0198	25	16	38.2	2.39
Vng0200	25	17	43.3	2.55
Vng0204	24	15	41.1	2.74
Vng0222	42	26	48	1.85
Vng0262	21	14	21.3	1.52
Vng0483	22	11	26.8	2.44
Vng0539	26	18	53.5	2.97
Vng0953	27	23	43.6	1.90
Vng01355	28	18	44.7	2.48
Vng1372	23	12	31.2	2.60
Vng1427	35	27	44.6	1.65
Vng1598	24	9	21.6	2.40
Vng1657	30	22	54.9	2.50
Vng1787	19	8	22.6	2.83
Vng1953	35	21	41.5	1.98
Vng2098	26	14	37.1	2.65
Vng2148	24	15	38.2	2.55
Vng2149	22	16	47.7	2.98
Vng2236	32	16	41.9	2.62
Vng2461	21	10	24.2	2.42
Vng2567	27	16	41.8	2.61
Vng2589	26	17	43	2.53
Vng6208	27	17	41.1	2.42
Vng6209	23	16	40.4	2.53
Average	26.86	16.86	38.64	2.35

Table B.1 Data for the analysis of the H-domain from the individual Sec substrates of *Halobacterium* sp NRC-1 (Table 4.1). The hydrophobicity of the H-domains was calculated using the Kyte-Doolittle scale. The average hydrophobicity of the amino acids in the H-domain was also calculated by dividing the hydrophobicity of the H-domain by the length of the H-domain.

Protein	Signal peptide length	Length of H domain	Hydrophobicity of H domain	Average Hydrophobicity
Aph	31	12	17.6	1.47
Chi	28	11	21.2	1.93
DmsA	32	13	18.6	1.43
HepB	39	21	23.6	1.12
HepD	38	19	22.5	1.18
Hly	30	16	24.4	1.53
Sub	28	10	15.7	1.57
Vng0199	34	16	39.3	2.46
Vng0615	25	14	24.2	1.73
Vng0818	28	9	22.7	2.52
Vng0819	27	15	21.1	1.40
Vng1036	19	9	18.2	2.02
Vng1268	22	9	19.4	2.16
Vng1952	35	17	42.6	2.51
Vng6162	32	15	16.9	1.13
Vng6266	33	13	16.0	1.23
Vng6296	23	10	13.2	1.32
Vng6323	24	12	21.7	1.81
Average	29.33	13.39	22.16	1.70

Table B.2 Data for the analysis of the H-domain from the individual Tat substrates of *Halobacterium* sp NRC-1 (Table 4.1). The hydrophobicity of the H-domains was calculated using the Kyte-Doolittle scale. The average hydrophobicity of the amino acids in the H-domain was also calculated by dividing the hydrophobicity of the H-domain by the length of the H-domain.

Protein	Signal peptide length	Length of H domain	Hydrophobicity of H domain	Average Hydrophobicity
ArtI	19	12	31.6	2.63
ArgT	22	10	28.3	2.83
BglX	20	11	24.6	2.24
CysP	25	11	27.4	2.49
DegP	26	16	29.4	1.84
DegQ	27	15	28.9	1.93
DsbA	19	12	30.8	2.57
FadL	27	10	26.9	2.69
GlnH	22	12	28.3	2.36
HdeA	24	15	41.3	2.75
LolA	23	12	25.7	2.14
MalE	24	12	23.9	1.99
MdoG	22	10	18.7	1.87
MepA	19	9	26.9	2.99
OmpA	21	12	30.2	2.52
OmpC	21	13	32.4	2.49
OmpT	20	12	27.5	2.29
OmpX	23	15	29.8	1.99
PhoA	21	12	25.7	2.14
PhoE	21	13	30.1	2.32
RbsB	25	14	28.2	2.01
Zra	26	13	25.4	1.95
Average	22.59	12.32	28.27	2.32

Table B.3 Data for the analysis of the H-domain from the individual Sec substrates of *E. coli* (Table 4.1). The hydrophobicity of the H-domains was calculated using the Kyte-Doolittle scale. The average hydrophobicity of the amino acids in the H-domain was also calculated by dividing the hydrophobicity of the H-domain by the length of the H-domain.

Protein	Signal peptide length	Length of H domain	Hydrophobicity of H domain	Average Hydrophobicity
AmiA	34	14	20.5	1.46
BisZ	37	20	25.0	1.25
DmsA	45	21	25.8	1.23
FdnG	33	20	27.4	1.37
FdoG	33	20	25.0	1.25
FhuD	30	12	20.5	1.71
HyaA	45	19	20.9	1.10
HybA	26	15	22.2	1.48
HybO	37	14	22.7	1.62
NapA	31	21	26.7	1.27
NapG	41	21	27.2	1.30
NrfC	27	13	23.9	1.84
PcoA	32	16	18.7	1.17
SufI	27	16	26.6	1.66
TorA	42	19	27.1	1.43
WcaM	28	11	20.1	1.83
YacK	28	16	23.0	1.44
YagT	48	16	22.5	1.41
YcdB	35	15	24.7	1.65
YedY	44	15	21.1	1.41
YnfE	42	19	25.7	1.35
YnfF	45	19	22.0	1.16
Average	35.91	16.91	23.60	1.43

Table B.4 Data for the analysis of the H-domain from the individual Tat substrates of *E. coli* (Table 4.1). The hydrophobicity of the H-domains was calculated using the Kyte-Doolittle scale. The average hydrophobicity of the amino acids in the H-domain was also calculated by dividing the hydrophobicity of the H-domain by the length of the H-domain.

	Activity	Activity – background proteases	Average	Standard deviation (±)
SptA +His 1	0.23	0.11	0.105	0.012
SptA +His 2	0.23	0.11		
SptA +His 3	0.21	0.09		
SptA +His 4	0.23	0.11		
SptA - His	0.32	0.20		
Background	0.12			

Table B.5 Data for the activity *H. volcanii* H26 transformed with pSPTAHIS (Table 4.4). Activity of the supernatant was measured using the azocasein assay. Background protease activity was measured from untransformed *H. volcanii* H26 was then subtracted from the other results

	Experiment 1	Experiment 2	Activity	Standard deviation (±)
H.hispanica Supernatant	0.160	0.135	0.342	0.018
H. volcanii + pAMY Supernatant	0.196	0.231	0.225	0.025
Blank	0.454	0.424		

Table B.6 Data for the activity of AmyH secreted by *H. hispanica* and *H. volcanii* transformed with pAMYH. Experiment 1 and 2 are duplicate experiments and absorbance was measured at OD_{620}. Activity is the average of the two experiments and the blank has also been taken into account.

Appendix C: Supplementary results from Chapter 5

Found below is a data table relevant to the work reported in Chapter 5.

	Activity	Activity – background proteases
Monensin (1.5µM/ml)	0.442	0.100
CCCP (0.5µM/ml)	0.356	0.014
Positive control	0.458	0.116
Background	0.342	

Table C.1 Data for *H. volcanii* H26 transformed with pSPTAHIS and cultured with the ionophores, Monensin or CCCP (Table 5.2). SptA activity was measured in the supernatant using the azocasein assay. Background protease activity was measured from untransformed *H. volcanii* H26 and then subtracted from the other results.

Appendix D: Supplementary results from Chapter 6

Found below are a number of data tables relevant to the work reported in Chapter 7 which are represented by graphs within the main body of text.

Temp ºC	Experiment 1		Experiment 2	
	NaCl	KCl	NaCl	KCl
55	100	100	100	100
55.3	95.19	97.60		
56.5	95.65	99.88	103.44	122.38
58.3	98.17	68.02		
60.6	94.39	60.72	96.66	81.62
63.2	84.90	24.79		
65.9	81.24	2.51	93.08	6.15
68.5	60.29	0		
71	56.06	0	49.43	0.08
73.1	18.76	0		
74.6	14.30	0	9.06	0.00
75.4	9.15	0		

Table D.1 Data for the stability of AmyH when incubated at varying temperatures for 20 minutes (Figure 6.2). Experiment 1 is the data used for the graph. Experiment 2 was also performed with fewer temperature points although the data is not mentioned it reveals similar results. Absorbances were measured at OD_{620} and then activity was converted to percentages with time 55ºC used as 100%.

Time (Hours)	Experiment 1		Experiment 2		Percentage	
	NaCl	KCl	NaCl	KCl	NaCl	KCl
0	0.976	0.904	0.983	0.899	100	100
1	0.921	0.872	0.926	0.858	77.28	80.09
2	0.946	0.891	0.945	0.882	79.12	82.08
3	0.904	0.867	0.883	0.846	74.76	79.3
4	0.912	0.876	0.896	0.853	75.65	80.05
5	0.929	0.882	0.918	0.791	77.28	77.45

Table D.2 Data for the stability of AmyH when incubated at 50ºC for 5 hours (Figure 6.3). Experiment 1 and 2 are duplicate experiments with absorbance measured at OD_{620}. The average activity was then determined as a percentage with time 0 used as 100%.

Time (Mins)	Experiment 1		Experiment 2		Percentage	
	NaCl	KCl	NaCl	KCl	NaCl	KCl
0	1.280	1.177	1.544	1.271	100.00	100.00
5	1.475	1.28	1.387	1.233	101.35	102.66
10	1.555	1.344	1.580	1.553	111.01	118.34
15	1.477	1.388	1.463	1.267	104.11	108.46
20	1.386	0.826	1.306	0.826	95.34	67.48
25	1.476	0.673	1.228	0.837	95.75	61.68
30	1.339	0.285	1.286	0.432	92.95	29.29
35	1.218	0.242	1.268	0.250	88.03	20.10
40	1.204	0.013	1.254	0.001	87.04	0.57
45	1.252	0.000	1.209	0.000	87.15	0.00
50	1.198	0.000	1.210	0.000	85.27	0.00
55	1.211	0.000	1.226	0.000	86.30	0.00
60	1.025	0.000	0.996	0.000	71.57	0.00

Table D.3 Data for the stability of AmyH when incubated at 70ºC for 1 hour (Figure 6.4). Experiment 1 and 2 are duplicate experiments and absorbance was measured at OD_{620}. The average activity was then determined as a percentage with time 0 used as 100%.

Time (Minutes)	Experiment 1	Experiment 2	Average
1	0.196	0.118	0.157
5	0.270	0.225	0.248
10	0.345	0.364	0.355
15	0.344	0.275	0.310
30	0.221	0.140	0.181
60	0.092	0.062	0.077
120	0.046	0.0093	0.070

Table D.4 Data for the activity of AmyH when allowed to fold for varying amounts of time (Figure 6.5). Experiment 1 and 2 are duplicate experiments and absorbance was measured at OD_{620}.

	Experiment 1	Experiment 2	Average
+ $CaCl_2$	0.112	0.109	0.111
- $CaCl_2$	0.108	0.109	0.109

Table D.5 Data for the activity of AmyH when allowed to fold in the presence/absence of $CaCl_2$ (Table 6.1). Experiment 1 and 2 are duplicate experiments and absorbance was measured at OD_{620}.

DTT concentration	Experiment 1	Experiment 2	Average
0	0.048	0.027	0.038
1	0.087	0.158	0.123
2	0.123	0.131	0.127
3	0.130	0.137	0.134
4	0.167	0.155	0.161
5	0.155	0.175	0.165
10	0.198	0.232	0.215
15	0.197	0.195	0.196
20	0.174	0.189	0.182
30	0.10	0.115	0.108

Table D.6 Data for the activity of AmyH when allowed to fold in varying concentrations of DTT (Figure 6.6). Experiment 1 and 2 are duplicate experiments and absorbance was measured at OD_{620}.

pH	Experiment 1	Experiment 2	Average
2.5	0.018	0.008	0.013
3.0	0.000	0.041	0.021
3.5	0.000	0.030	0.015
4.0	0.008	0.060	0.034
4.5	0.064	0.078	0.071
5.0	0.06	0.095	0.078
5.5	0.097	0.106	0.102
6.0	0.127	0.106	0.117
6.5	0.121	0.141	0.131
7.0	0.205	0.206	0.206
7.5	0.287	0.362	0.325
8.0	0.333	0.360	0.347
8.5	0.286	0.290	0.288
9.0	0.216	0.224	0.220

Table D.7 Data for the activity of AmyH when allowed to fold at varying pH (Figure 6.7). Experiment 1 and 2 are duplicate experiments and absorbance was measured at OD_{620}.

Salt	Experiment 1		Experiment 2		Average	
Conc.(M)	NaCl	KCl	NaCl	KCl	NaCl	KCl
0.5	0.071	0.041	0.042	0.052	0.057	0.047
1.0	0.094	0.058	0.095	0.072	0.095	0.065
1.5	0.115	0.091	0.118	0.096	0.117	0.094
2.0	0.166	0.13	0.143	0.127	0.155	0.129
2.5	0.165	0.153	0.198	0.148	0.182	0.151
3.0	0.175	0.169	0.177	0.148	0.176	0.159
3.5	0.185	0.174	0.178	0.168	0.182	0.171
4.0	0.172	0.173	0.159	0.177	0.166	0.175
4.5	0.152	0.173	0.156	0.139	0.154	0.156

Table D.8 Data for the activity of AmyH when allowed to fold in varying salt concentrations of NaCl and KCl (Figure 6.8). Experiment 1 and 2 are duplicate experiments and absorbance was measured at OD_{620}.

Appendix E: Published work

the FEBS Journal

Bioenergetic requirements of a Tat-dependent substrate in the halophilic archaeon *Haloarcula hispanica*

Daniel C. Kwan[1], Judith R. Thomas[2,*] and Albert Bolhuis[1]

1 Department of Pharmacy and Pharmacology, University of Bath, UK
2 Department of Biological Sciences, University of Warwick, Coventry, UK

Keywords
halophilic archaea; protein translocation; signal peptide; sodium motive force; twin-arginine translocase

Correspondence
A. Bolhuis, Department of Pharmacy and Pharmacology, University of Bath, Bath BA2 7AY, UK
Fax: +44 1225 386114
Tel: +44 1225 383813
E-mail: a.bolhuis@bath.ac.uk

***Present address**
Systems Biology Laboratory UK, Abingdon, UK

(Received 6 May 2008, revised 1 October 2008, accepted 13 October 2008)

doi:10.1111/j.1742-4658.2008.06740.x

Twin-arginine translocase (Tat) is involved in the translocation of fully folded proteins in a process that is driven by the proton motive force. In most prokaryotes, the Tat system transports only a small proportion of secretory proteins, and Tat substrates are often cofactor-containing proteins that require folding before translocation. A notable exception is found in halophilic archaea (haloarchaea), which are predicted to secrete the majority of their proteins through the Tat pathway. In this study, we have analysed the translocation of a secretory protein (AmyH) from the haloarchaeon *Haloarcula hispanica*. Using both *in vivo* and *in vitro* translocation assays, we demonstrate that AmyH transport is Tat-dependent, and, surprisingly, that its secretion does not depend on the proton motive force but requires the sodium motive force instead.

The twin-arginine translocation (Tat) pathway is a system for protein translocation that is found in the cytoplasmic membrane of most prokaryotes and in the thylakoid membrane of chloroplasts [1]. The Tat system usually requires two or three membrane-bound components, denoted TatA, TatB and TatC. TatA and TatB are similar in sequence and structure and contain one membrane-spanning domain, whereas TatC contains six membrane-spanning domains. All three proteins have distinct functions, although many organisms (including most Gram-positive bacteria and archaea) seem to lack TatB-like proteins [1]. The Tat system has the unique ability to translocate fully folded proteins. This is in stark contrast to the Sec machinery, the main system for protein translocation in prokaryotes, which is only able to translocate proteins that are in an unfolded state [2]. In prokaryotes, many Tat substrates bind complex cofactors that are incorporated in the cytoplasm [3], which explains the need for a system that is able to translocate folded proteins. There are, however, also Tat-dependent substrates that do not bind cofactors, and it may be that these require the Tat system simply because they fold very rapidly. The latter may be the reason why the Tat system appears to play a dominant role in protein translocation in halophilic archaea (haloarchaea) [4,5]. These organisms live in concentrated brine, with the main salt usually being NaCl. To deal with the osmotic stress, haloarchaea have adapted a 'salt-in' strategy, and the intracellular concentration of

Abbreviations
AmyH, α-amylase from *H. hispanica* (AmyH); CCCP, carbonyl cyanide *m*-chlorophenylhydrazone; HAP, hemagglutinin protease; IMVs, inverted membrane vesicles; MIC, minimal inhibitory concentration; PMF, proton motive force; SMF, sodium motive force; Tat, twin-arginine translocase.

salt (predominately KCl) is equal to the extracellular salt concentration [6]. It has been suggested that proteins fold very rapidly under these conditions due to salting-out effects [4]. From this, it follows that many secretory proteins in haloarchaea fold before translocation and thus require the Tat system for export. Genomic surveys have indeed shown that at least 60–70% of the secretory proteins in halophilic archaea contain a signal peptide with a characteristic twin-arginine motif, while other organisms usually secrete most of their proteins (> 90%) through the Sec pathway [4,7]. The dominant role of the Tat system in haloarchaea was corroborated by the observation that the Tat system is essential for viability in these organisms [8,9].

The Tat system in bacteria and chloroplasts is driven by the proton motive force (PMF). It was first identified in chloroplasts as a protein translocation system that relied on the pH gradient across the thylakoid membrane [10], and is therefore sometimes also called the ΔpH pathway. More recent data have shown that, in thylakoids, the electrical gradient Δψ can also contribute to Tat-dependent translocation [11], although it should be noted that the Δψ normally forms only a small part of the PMF in thylakoids. In bacteria, involvement of the PMF was first shown through inhibition of translocation of the precursor of the *Escherichia coli* Tat substrate TorA (preTorA) by the protonophore carbonyl cyanide *m*-chlorophenylhydrazone (CCCP) [12]. Recently it has been shown that translocation of another *E. coli* Tat substrate, preSufI, is independent of the ΔpH and only requires Δψ for export [13]. Here, we report the development of an *in vitro* assay for Tat-dependent translocation in the haloarchaeon *Haloarcula hispanica*. Using this *in vitro* assay, as well as *in vivo* translocation assays, we show that secretion of a Tat-dependent α-amylase does not depend on the PMF but is driven by the sodium motive force (SMF) instead.

Results

AmyH is a Tat-dependent substrate

We have previously reported that the α-amylase from *H. hispanica* (AmyH) is probably a Tat-dependent substrate as (a) the signal peptide contains a characteristic twin-arginine motif, and (b) the precursor of AmyH (preAmyH) in the cytoplasm is fully active, indicating that it folds before translocation [14]. To provide further evidence for its Tat dependency, the *amyH* gene was cloned in a haloarchaeal expression vector and the two codons encoding the two arginine residues in the Tat motif (positions 14 and 15 in the signal peptide) were altered to change the arginines into lysines. Next, plasmids encoding preAmyH and the signal peptide mutant (denoted preAmyH-KK) were used to transform *Haloferax volcanii*, a haloarchaeon that lacks endogenous amylase activity. The secretion of AmyH was monitored on agar plates containing starch. As shown in Fig. 1A, *H. volcanii* expressing wild-type preAmyH secreted significant amounts of amylase activity into the medium, whereas the strain producing preAmyH-KK produced only a very small halo on the starch plates. These results were confirmed by western blotting. As shown in Fig. 1B, wild-type AmyH was exported in *H. volcanii*, but the amount of preAmyH-KK was very low (Fig. 1B, compare lanes 2 and 4). A small amount of preAmyH-KK appears to be present

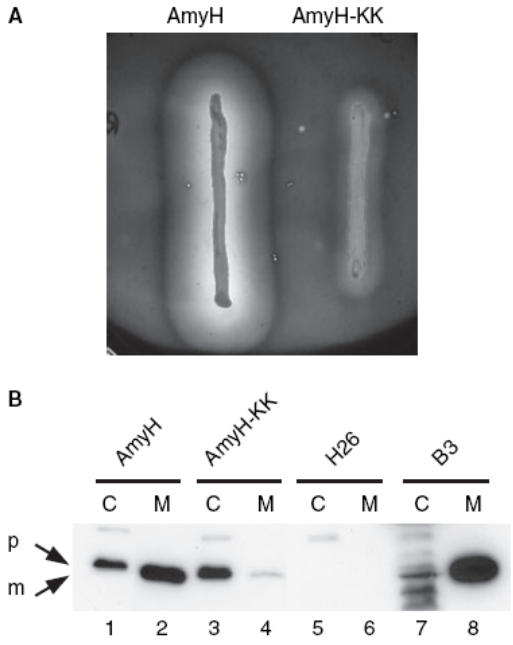

Fig. 1. AmyH is not secreted when the double arginine in the signal peptide is changed into a double lysine. (A) *H. volcanii* transformed with plasmid pSY-AmyH (encoding AmyH) or pSY-AmyH-KK (encoding AmyH-KK) were grown on rich medium agar plates containing 0.5% starch. Plates were then stained with iodine solution (2% KI, 0.2% I_2). A clear halo, which is an indication of starch degradation by AmyH released into the medium, is only seen around cells producing wild-type AmyH. (B) Cells were grown in liquid medium, and cells (C) and medium (M) were separated by centrifugation. AmyH was visualized by SDS–PAGE and western blotting using AmyH-specific antibodies. Lanes 1 and 2, *H. volcanii* producing AmyH; lanes 3 and 4, *H. volcanii* producing AmyH-KK; lanes 5 and 6, *H. volcanii* lacking AmyH; lanes 7 and 8, *H. hispanica* B3 AmyH-overproducing mutant. p, precursor; m, mature AmyH.

in the medium, but we cannot exclude the possibility that this is the result of cellular lysis, particularly as the precursor and mature forms of AmyH are poorly separated on SDS–PAGE gels. In any case, it is obvious that changing the double arginine in the signal peptide to a double lysine severely affects translocation of preAmyH, demonstrating that this protein is a Tat-dependent substrate. Figure 1B also shows two additional controls – *H. volcanii* H26, which does not contain the *amyH* gene (demonstrating that *H. volcanii* does not produce another protein recognized by the AmyH antibodies), and *H. hispanica* B3, which is an AmyH-overproducing mutant of the native *H. hispanica* strain [14].

Effect of ionophores on AmyH secretion

Ionophores can be used to disrupt various gradients across membranes, and they are therefore useful in analysis of the bioenergetics of cellular processes in prokaryotes. To investigate the effect of ionophores on the secretion of AmyH, we first measured the minimal inhibitory concentration (MIC) of several ionophores, and then monitored the effect on amylase secretion in *H. hispanica* B3 at 50% of the MIC (Table 1). Three of the ionophores chosen are frequently used to determine the effect of the proton motive force in prokaryotic protein translocation. These are carbonyl cyanide m-chlorophenylhydrazone (CCCP), which is a proton carrier and uncoupler that disrupts the entire proton motive force; valinomycin, a K^+-specific ionophore that dissipates the $\Delta\psi$; and nigericin, a K^+/H^+ antiporter that dissipates the ΔpH. Two other ionophores used are monensin, which is similar to nigericin but with a high specificity for Na^+ ions, and nonactin, which is similar to valinomycin but also shows some affinity to other ions such as Na^+ and NH_4^+ (although its highest affinity is for K^+).

The MIC values of *H. hispanica* cells for these ionophores differed greatly, varying from 0.0625 μM for nigericin to 40 μM for valinomycin (Table 1). When *H. hispanica* cells were grown in the presence of ionophores at a concentration of 50% of the MIC, AmyH was secreted at normal levels in the presence of all ionophores with the exception of monensin (Table 1). In particular, the lack of effect of CCCP is remarkable as it affects both the electrical and chemical components of the proton motive force. Valinomycin, nigericin and nonactin also did not affect AmyH secretion, suggesting that AmyH secretion is independent of the PMF. In contrast, cells grown in the presence of sub-MIC concentrations of monensin do not secrete detectable amounts of AmyH. Monensin is a sodium/proton antiporter that has been used as a tool in a number of organisms to demonstrate involvement of the sodium motive force (SMF) in cellular processes [15–17]. The lack of secretion of AmyH in the presence of monensin suggests that export of AmyH might depend on the SMF.

To study the effect of monensin in more detail, pulse–chase experiments were performed in which *H. hispanica* B3 cells were radiolabelled for 5 min with ^{35}S-methionine (pulse), after which an excess of 'cold' methionine was added (chase). As shown in Fig. 2, in the absence of ionophore, 74% of AmyH is in the mature processed form after 10 min, and 85% of AmyH is mature after 30 min of chase treatment. The rate of translocation

Table 1. Minimal inhibitory concentrations of ionophores and their effects on AmyH secretion.

Ionophore	MIC (μM)	AmyH secretion
CCCP	0.6	+
Monensin	2.5	–
Nigericin	0.0625	+
Nonactin	20	+
Valinomycin	40	+

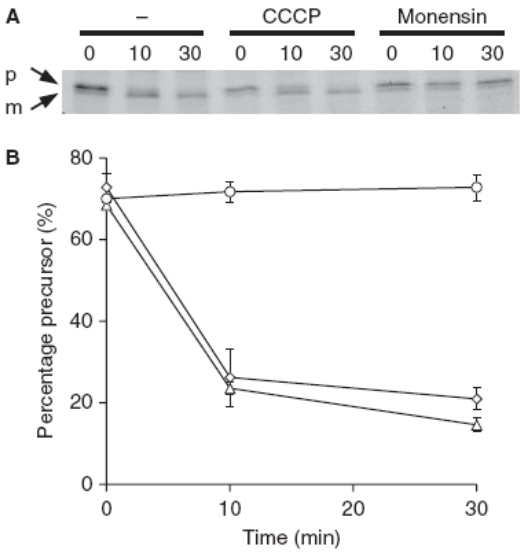

Fig. 2. Effect of ionophores on translocation of AmyH. (A) Pulse–chase reactions were performed in the absence or presence of CCCP or monensin. Samples were taken after 0, 10 and 30 min of chase as indicated. p, precursor; m, mature AmyH. (B) The kinetics of processing were plotted as the percentage of AmyH still in the precursor form at the time of sampling. The error bars shown were calculated from two independent pulse–chase experiments. Triangles, no addition; diamonds, with CCCP; circles, with monensin.

measured here is somewhat faster then reported previously [14], which is probably because we optimized the pulse–chase protocol for *H. hispanica*.

Treatment of cells shortly before a pulse–chase experiment confirmed that monensin does indeed block the translocation of AmyH. In the presence of 5 μM monensin, the precursor is not converted into the mature form, and the precursor/mature ratio remains constant over a period of 30 min. This shows that precursor processing, which occurs during or shortly after translocation on the *trans* side of the membrane, is almost completely blocked. On the other hand, translocation of AmyH was not affected at all by the addition of 50 μM CCCP, despite the fact that the concentration used was more than 80 times the MIC value. Thus, even concentrations of CCCP that completely stop growth were not sufficient to block or slow down precursor processing during the time period of the pulse–chase reactions, which is another clear indication that AmyH secretion does not depend on the proton gradient.

In vitro translocation

It is conceivable that the effect of monensin on AmyH secretion is not directly due to dissipation of the SMF. We have, for instance, observed that addition of high concentrations of monensin (50 μM) leads to cell lysis within a few minutes, suggesting that secondary effects may play a role. It was therefore important to investigate the role of the sodium gradient using an experimental set-up that does not require ionophores. For that purpose, we sought to develop an *in vitro* translocation assay. The basic principle of this assay is to synthesize radiolabelled preAmyH *in vitro* and import it into inverted membrane vesicles (IMVs). A cell-free protein synthesis system for haloarchaea has been developed [18], but is unfortunately not very efficient. We therefore chose to use a commercially available system, but, because it would only work under low-salt conditions, it was important to establish whether *in vitro* synthesized AmyH could fold into its native conformation. We have previously shown that purified AmyH unfolds in the presence of urea and low salt concentrations [14]. Various conditions for refolding were tested, and it appeared that reducing conditions (> 5 mM dithiothreitol) were essential for refolding; even in the absence of salt, AmyH refolded with reasonable efficiency (approximately 60%), and this efficiency increased with higher concentrations of salt (data not shown). As the *E. coli* transcription/translation system we used is under reducing conditions and contains approximately 60 mM KCl, it seems likely that AmyH synthesized in such a system is able to fold correctly. In low salt, AmyH has a somewhat loose structure that becomes more tightly folded upon the addition of salt [14]. It was therefore anticipated that, if correctly folded, *in vitro* synthesized AmyH would be much more resistant to protease degradation in high salt compared to low salt. This was observed (data not shown), indicating that the *in vitro* synthesized AmyH is folded in its correct conformation.

For the *in vitro* translocation system, we first established conditions under which we could detect translocation of *in vitro* synthesized preAmyH into IMVs. The goal was to mimic the conditions found in haloarchaea, i.e. a high concentration of NaCl in the extracellular milieu and an equimolar concentration of KCl in the cytoplasm. For our IMV-based system, these conditions are reflected by a high concentration of NaCl inside the vesicles, and a similarly high concentration of KCl outside the vesicles. We therefore first prepared radiolabelled preAmyH that was synthesized in an *E. coli* based cell-free translation system, which was then dialysed against a buffer containing 2.5 M KCl. A concentrated stock of IMVs was prepared in a buffer containing 2.5 M NaCl. As shown in Fig. 3A,

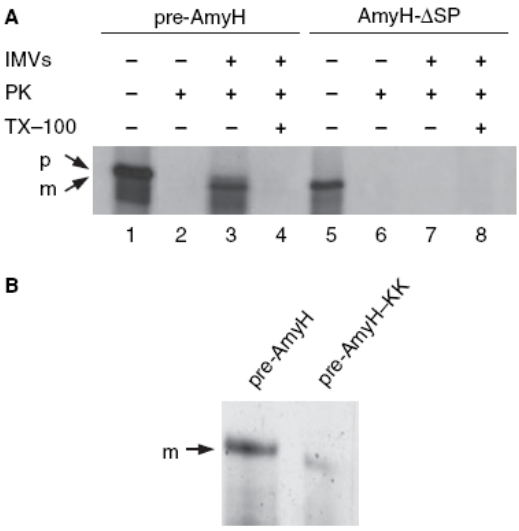

Fig. 3. *In vitro* translocation assay of preAmyH. (A) Lanes 1–4 and 5–8 show preAmyH and AmyH-ΔSP, respectively. PreAmyH and AmyH-ΔSP were incubated in the presence or absence of IMVs, proteinase K (PK) and/or Triton X-100 (TX-100) as indicated. The loading for the translation reactions in lanes 1 and 5 is 5% of the amount used in the translocation assays in lanes 2–4 and 6–8. (B) *In vitro* translocation was performed as in lane 3 in (A) using either *in vitro* synthesized preAmyH or preAmyH-KK in which the twin arginines were altered to two lysines.

preAmyH could be synthesized efficiently *in vitro* (lane 1); the same was found for a mutant lacking most of its signal peptide (denoted AmyH-ΔSP; it contains only the first two residues of the signal peptide; lane 5). When *in vitro* synthesized preAmyH was incubated in the presence of 20-fold diluted IMVs that were energized by addition of ATP and NADH, a protease-protected band could be observed that was not seen in the absence of vesicles (Fig. 3A, compare lanes 2 and 3). The protease-protected band was slightly smaller than full-length preAmyH, indicating processing of the signal peptide. As expected, AmyH was fully degraded after import when the IMVs were solubilized by addition of the detergent Triton X-100 (lane 4). When AmyH-ΔSP was incubated in the presence of IMVs, no protease-protected band could be observed (compare lanes 3 and 7). Thus the protease-protected band is only observed in the presence of a signal peptide, demonstrating that we have developed a genuine *in vitro* translocation system for Tat-dependent translocation in *H. hispanica*.

To verify that the *in vitro* translocation observed was a Tat-dependent process, the translocation assay was also performed with *in vitro* synthesized preAmyH-KK. As shown in Fig. 3B, significantly less AmyH-KK was protected from protease degradation, demonstrating that the observed *in vitro* translocation is a Tat-dependent process.

The next step was to investigate the bioenergetics of the haloarchaeal Tat system using the *in vitro* translocation system. For this purpose, experiments as above were repeated in the presence and absence of ATP/NADH, and reactions were performed using *in vitro* synthesized preAmyH that was either dialysed against KCl-containing buffer or NaCl-containing buffer. In the latter case, there was no sodium gradient, as the concentrations of NaCl inside and outside the IMVs were identical. As shown in Fig. 4, whether ATP and NADH were present or not only resulted in a fairly small difference in the efficiency of translocation; the translocation efficiency in the absence of ATP/NADH was approximately 70% of that in the presence of ATP/NADH (compare lanes 3 and 4). A much more significant fivefold reduction in efficiency was seen in the absence of a sodium gradient (compare lanes 3 and 6), under which conditions only a small fraction of preAmyH was translocated. This was even further reduced in the absence of ATP and NADH (lane 7).

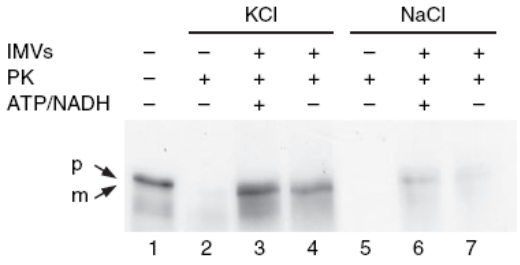

Fig. 4. *In vitro* translocation of preAmyH in the presence and absence of a sodium gradient. Translocation reactions were performed in the presence or absence of IMVs, proteinase K (PK) and/or ATP plus NADH as indicated. Lane 1 contains the translation reaction of preAmyH, reactions in lanes 2–4 were performed using preAmyH dialysed against a buffer containing 2.5 M KCl, and reactions in lanes 5–7 were performed using preAmyH dialysed against a buffer containing 2.5 M NaCl.

Discussion

In the present study, we show that the α-amylase AmyH from *H. hispanica* is a Tat-dependent protein, the translocation of which depends on the SMF. Its Tat dependence was expected, as the signal peptide of AmyH contains a characteristic twin-arginine motif. We had also shown previously that preAmyH in the cytoplasm is fully active, indicating that it folds before translocation [14]. Here we show that changing the double arginine to a double lysine blocks translocation both *in vivo* and *in vitro*; such a mutation does not normally affect a Sec substrate, and indeed similar RR to KK mutations have been produced to show the Tat dependency of the α-amylase from the haloarchaeon *Natronococcus* sp. strain Ah36, for example [5]. In bacteria such as *E. coli* or *Bacillus subtilis*, involvement of the Tat system in export has also been shown through deletion of Tat components [19–21]; however, the Tat system is essential in haloarchaea and cannot be deleted [8,9]. Our observation that AmyH can only refold under reducing conditions further corroborates the Tat dependency of the protein, as the extracellular environment in which organisms such as *H. hispanica* thrive (shallow salt lakes and solar salterns) is probably mostly oxidizing. Thus, AmyH would not be able to fold efficiently at the *trans* side of the membrane, and seems to require the more reducing environment of the cytoplasm to become active. The reason that AmyH cannot fold under oxidizing conditions may be due to the presence of the cysteine residues in the protein. These probably do not form a disulfide bond, but, under oxidizing conditions, it seems likely that if the protein is (not yet) folded, intra- or intermolecular disulfide bonds will be readily formed, leading to incorrect folding and thus an inactive protein. We presume that, once AmyH is folded in its correct conformation, the protein remains stable in the more oxidizing environment into which it is secreted.

The most interesting finding of our study is that AmyH secretion is independent of the PMF, and appears to depend on the SMF instead. In *E. coli*, and most likely also in other bacteria, the Tat system depends on the PMF [12]. Here we show involvement of the SMF in *H. hispanica in vivo*, as translocation of preAmyH was only affected by the sodium ionophore monensin. Just as significant was our observation that AmyH secretion was not affected by the ionophores CCCP, valinomycin, nigericin or nonactin, clearly indicating that the proton gradient is not involved. However, we could not exclude the possibility of indirect effects of monensin on AmyH translocation, and it was therefore important to develop an experimental system that did not require the use of ionophores. As shown using an *in vitro* translocation system, transport of preAmyH did not depend on the presence or absence of ATP and NADH, although the efficiency was somewhat increased when ATP and NADH were present. It is not clear whether the observed differences in the presence or absence of ATP/NADH were significant, but it is conceivable that the SMF is maintained more stably in vesicles in the presence of ATP and NADH; in most haloarchaea, the main source of energy for the extrusion of sodium and accumulation of potassium is the PMF, which in turn can be generated by the respiratory chain (at the expense of NADH) or by ATP synthase (at the expense of ATP) [22]. We did, however, observe some translocation in the absence of a sodium gradient when ATP and NADH were present. This might suggest that the PMF could drive Tat-dependent translocation in *H. hispanica*, albeit very inefficiently.

To our knowledge, involvement of the SMF in protein transport has only been shown in *Vibrio* species. Secretion of a Sec substrate, hemagglutinin protease (HAP) in *Vibrio cholerae*, is strongly affected by treatment of cells with monensin, but is hardly affected by CCCP [23]. Using IMVs isolated from a Na^+ pump-deficient mutant, the Sec pathway of *Vibrio alginolyticus* was also shown to be stimulated by the sodium gradient [24]. In the latter case, a requirement for ATP was also demonstrated, but that was unsurprising as translocation of Sec substrates such as HAP depend on the ATPase SecA, which is a central component of the bacterial Sec machinery.

Other cellular processes have also been shown to be dependent on the SMF. The archaeon *Methanosarcina barkeri* requires the SMF for oxidation of methanol [15], while both the haloarchaeon *Halobacterium salinarum* (*halobium*) and the thermophilic bacterium *Bacillus* sp. TA2.A1 require the SMF for uptake of glutamate [16,25,26]. We have shown here for the first time that the SMF is required for secretion of a Tat-dependent substrate. Future studies are required to establish whether this sodium gradient is only used for specific proteins or by particular organisms, or whether the SMF is more generally used by all haloarchaea for Tat-dependent protein translocation. It is of interest to note that the genomes of all haloarchaea that have been sequenced to date contain a Tat component with a unique topology that is not found in other organisms [4,8]. This component, denoted TatC2 in *H. salinarum* or TatCt in *H. volcanii* [9], consists of a natural fusion of two TatC-like proteins. As TatC2 appears to be specific to haloarchaea, it is conceivable that it is required for adaptation of the Tat pathway to highly saline conditions. If all haloarchaea use the SMF for Tat-dependent translocation, it is tempting to speculate that TatC2 has a role in linking protein secretion to the sodium gradient.

Experimental procedures

Chemicals

All chemicals used were purchased from Sigma-Aldrich (Poole, UK) or Fisher Scientific (Loughborough, UK).

Strains and growth conditions

Wild-type *H. hispanica* and *H. hispanica* B3 have been described previously [14] and were routinely grown at 45 °C on rich medium containing 0.5% peptone (Oxoid, Basingstoke, UK), 0.1% yeast extract (Difco, Becton Dickinson, Oxford, UK), and 23% salt water (18.4% NaCl, 2.7% $MgSO_4{\cdot}7H_2O$, 2.3% $MgCl_2{\cdot}6H_2O$, 0.54% KCl and 0.056% $CaCl_2$). Minimal medium for *H. hispanica* contained 16% NaCl, 6.4% $MgCl_2{\cdot}6H_2O$, 0.64% K_2SO_4, 10 mM NH_4Cl, 2 mM $CaCl_2$, 0.5 mM K_2SO_4, 5 mM $NaHCO_3$, 0.2% glycerol, trace elements (0.36 $mg{\cdot}L^{-1}$ $MnCl_2{\cdot}4H_2O$, 0.44 $mg{\cdot}L^{-1}$ $ZnSO_4{\cdot}7H_2O$, 2.3 $mg{\cdot}L^{-1}$ $FeSO_4{\cdot}7H_2O$ and 0.05 $mg{\cdot}L^{-1}$ $CuSO_4{\cdot}5H_2O$) and vitamins (1 $mg{\cdot}L^{-1}$ thiamine and 0.1 $mg{\cdot}L^{-1}$ biotin). *Haloferax volcanii* H26 has been described previously [27] and was routinely grown at 45 °C in rich medium (Hv-YPC) containing 0.5% yeast extract, 0.1% peptone, 0.1% casamino acids and 18% salt water (14.4% NaCl, 2.1% $MgSO_4{\cdot}7H_2O$, 1.8% $MgCl_2{\cdot}6H_2O$, 0.42% KCl, 0.056% $CaCl_2$ and 12 mM Tris/HCl pH 7.5). Solid media were prepared by the addition of 1.5% agar. If required, mevinolin was added at 2 $\mu g{\cdot}mL^{-1}$ and novobiocin at 0.3 $\mu g{\cdot}mL^{-1}$.

E. coli was routinely grown in Luria–Bertani medium (0.5% yeast extract, 1% peptone, 1% NaCl); if required, 100 $\mu g{\cdot}mL^{-1}$ ampicillin was added. For construction of plasmids, *E. coli* JM109 (F′ *traD36 proA*$^+$*B*$^+$ *lacI*^{q} Δ*(lacZ) M15/*Δ*(lac-proAB) glnV44 e14*$^-$ *gyrA96 recA1 relA1*

endA1 thi hsdR17) was used. To prepare unmethylated DNA for efficient transformation of *H. volcanii*, *E. coli* ER2925 (New England Biolabs, Hitchin, UK) was used.

DNA techniques

Enzymes for restriction and ligation were purchased from Invitrogen. Transformation of *E. coli* and *H. volcanii* was performed as described previously [28,29].

PCR was performed using Dynazyme EXT (New England Biolabs) in the presence of 3% dimethylsulfoxide. The nucleotide sequences of primers used for PCR (5′ → 3′) are listed below; nucleotides identical to the template DNA are printed in capital letters and restriction sites used for cloning are underlined. All plasmids were verified by sequencing.

To construct pET-AmyH for use in *in vitro* transcription/translation (see below), *amyH* was amplified using chromosomal DNA of *H. hispanica* as template, and primers AmyH-T7a (atatcatATGAATCGACCCCGAATTACC GGCAG) and AmyH-T7b (atataagcttGTCTCCGTGGCG TGCCAGCTTACTG), and cloned into the *Nde*I and *Hin*dIII sites of plasmid pET21a (Novagen, Nottingham, UK). To construct pET-ΔSP-AmyH, *amyH* lacking most of the region encoding the signal peptide (residues 3–40) was amplified using primers AmyH-ΔSP-T7 (atatcatATGAATG TCGGCGATAGCGCGGTGTACCAG) and AmyH-T7b, and cloned into the *Nde*I and *Hin*dIII sites of plasmid pET21a. The Quickchange mutagenesis system (Stratagene, La Jolla, CA, USA) was used to construct pET-AmyH_KK, encoding AmyH in which the twin arginines of the signal peptide were mutated to twin lysines (AmyH-KK). The primers used for Quickchange mutagenesis were AmyKKfor (CCGGCAGTAAGCAGGCGTCT**aagaaa**ACC GTTCTGAAAGGAATCG) and AmyKKrev (GGCCGTC ATTCGTCCGCAGA**ttcttt**TGGCAAGACTTTCCTTAGC) (bold letters indicate the nucleotides encoding the mutated residues).

To construct pSY-AmyH, the *amyH* gene from *H. hispanica* was amplified using pET-AmyH as template, with primers AmyFor-*Nde*I (TTTGTTTAACTTTAAGAAGG AGATATACATATGAATCG) and AmyRev-*Nco*I (aaaac catGGGCTTTGTTAGCAGCCGGAT). The amplified fragment was ligated into the *Nde*I and *Nco*I sites of pSY1 [30]. A derivative (pSY-AmyH_KK) was also made from pSY-AmyH containing an *amyH* gene encoding AmyH in which the twin arginines of the signal peptide were mutated to twin lysines, using the Quickchange mutagenesis system and primers AmyKKfor and AmyKKrev as described above.

Western blotting

Proteins were separated by SDS–PAGE and immunoblotted on poly(vinylidene) difluoride membranes (Millipore, Watford, UK) using a semi-dry system. Amylase was visualized using specific antibodies and horseradish peroxidase anti-rabbit IgG conjugates (Promega, Southampton, UK) using the Pico West detection system (Perbio Science, Cramlington, UK).

Amylase activity assays

Amylase activity in buffer (50 mM Bistris pH 6.5, 4 M NaCl and 5 mM $CaCl_2$) was determined by measuring released reducing sugars, using the dinitrosalicylic acid method or the starch–iodine method as described previously [14].

Refolding of AmyH

AmyH was purified as described previously [14] and unfolded by dialysis against a buffer (50 mM Bistris pH 6.5) containing 6 M urea. AmyH was then refolded by rapid dilution (20-fold) in buffer (50 mM Bistris pH 6.5, 3.5 M NaCl, 5 mM $CaCl_2$ and 5 mM dithiothreitol). Activity was then measured using the starch–iodine method.

Minimal inhibitory concentrations of ionophores

Tubes containing 5 mL of rich medium containing various concentrations of ionophores were inoculated with 10^5 cells per mL *H. hispanica* B3 cells. The lowest concentration where no growth was observed after 48 h of growth was taken as the MIC. For incubation in the presence of sublethal concentrations of ionophores, cells were grown in the presence of the various ionophores at 50% of the MIC.

Pulse–chase protein labelling and immunoprecipitation

Cells of *H. hispanica* B3 were grown in rich medium until an attenuance at 660 nm of 0.6–0.8 was reached. Cells were collected by centrifugation (12 000 ***g*** for 2 min at room temperature), washed briefly in minimal medium, and then resuspended in minimal medium (attenuance at 660 nm of approximately 0.8). Cells were incubated for 1 h at 45 °C in a shaking incubator. Cells were pulsed for 5 min with 40 μCi [^{35}S]-methionine (Perkin Elmer, Waltham, MA, USA) per mL culture medium. Where indicated, 50 μM CCCP or 5 μM monensin was added at the end of the pulse period. Next, an excess of non-radioactive methionine was added (1 mg·mL^{-1}), and 1 mL samples were taken after 0, 10 and 30 min. Samples were immediately mixed with cold trichloroacetic acid (final concentration 15%), and kept on ice for at least 30 min. Cells and proteins were pelleted by centrifugation (20 800 ***g*** for 15 min at 4° C), and washed briefly twice with ice-cold acetone. Pellets were resuspended in 50 μL buffer (50 mM Tris/HCl pH 8, 1% SDS and 1 mM EDTA) and boiled for 10 min. Next, 1 mL Triton buffer (2% Triton X-100, 50 mM Tris/HCl pH 8, 150 mM

NaCl and 0.1 mM EDTA) was added, and insoluble precipitates were removed by centrifugation (20 800 ***g*** for 2 min at 4 °C). Samples were incubated for 2 h at room temperature in the presence of AmyH-specific polyclonal antibodies [14]. Next, 5 mg protein A–Sepharose washed in Triton buffer was added, and the samples were incubated for a further 2 h. The protein A–Sepharose beads were washed briefly three times with Triton buffer and boiled in 40 μL SDS–PAGE loading buffer. Samples were visualized using SDS–PAGE and a Fuji FLA-5000 phosphorimager (Fujifilm, Bedford, UK).

Isolation of inverted membrane vesicles

IMVs were essentially isolated as described previously [31]. In brief, *H. hispanica* cells were grown in rich medium until an attenuance at 660 nm of approximately 0.8 was reached. Cells were collected by centrifugation (6000 ***g*** for 20 min at 4 °C) and resuspended in buffer A (1.25 M NaCl, 50 mM Tris/HCl pH 7.5, 1 mM $CaCl_2$ and 25 mM $MgCl_2$) containing Complete protease inhibitor cocktail (Roche, Burgess Hill, UK). Cells were lysed by sonication, and cellular debris was removed by centrifugation for 10 min at 5000 ***g***. Next, membranes were collected by centrifugation for 30 min at 180 000 ***g*** in an Optima Max ultracentrifuge (Beckman, High Wycombe, UK), washed in buffer B (2.5 M NaCl, 50 mM Tris/HCl pH 7.5, 1 mM $CaCl_2$ and 25 mM $MgCl_2$), and finally resuspended in buffer B to a final protein concentration of 20 $mg \cdot mL^{-1}$. The membrane orientation was verified using the menadione-dependent NADH dehydrogenase activity assay [32]; with the method used, at least 75–80% of vesicles had an inside-out orientation.

In vitro translation

PreAmyH, preAmyH-KK and preAmyH-ΔSP were translated *in vitro* using the pET vectors described above and the *E. coli* T7 S30 extract system. Reactions were performed in the presence of [^{35}S]-methionine, according to the instructions of the manufacturer (Promega). After translation, reactions were dialysed against translocation buffer (2.5 M KCl, 50 mM Tris/HCl pH 7.5, 1 mM $CaCl_2$, 25 mM $MgCl_2$ and 5 mM dithiothreitol).

In vitro translocation reactions

In vitro translocation reactions were performed in 100 μL translocation buffer containing *in vitro* synthesized preAmyH, 3 mM Mg-ATP, 5 mM NADH and 5 μL IMVs. Reactions were incubated for 60 min at 45 °C. Next, proteinase K (0.5 $mg \cdot mL^{-1}$) was added, and samples were incubated for 60 min at 37 °C. Reactions were stopped by the addition of four volumes of 25% trichloroacetic acid, and, after 30 min on ice, the proteins were pelleted by centrifugation (20 800 ***g*** for 15 min at 4 °C). Pellets were washed with ice-cold acetone, dried in air, and resuspended in SDS–PAGE loading buffer. Samples were analysed by SDS–PAGE and fluorography.

Acknowledgements

We thank Dr Xiao-Feng Tang (College of Life Sciences, Wuhan University, China) for providing plasmid pSY1. D.C.K. and J.R.T. were supported by the Biotechnology and Biological Sciences Research Council, and A.B. is the recipient of a Royal Society University Research Fellowship.

References

1 Robinson C & Bolhuis A (2004) Tat-dependent protein targeting in prokaryotes and chloroplasts. *Biochim Biophys Acta* **1694**, 135–147.
2 Gold VA, Duong F & Collinson I (2007) Structure and function of the bacterial Sec translocon. *Mol Membr Biol* **24**, 387–394.
3 Palmer T, Sargent F & Berks BC (2005) Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol* **13**, 175–180.
4 Bolhuis A (2002) Protein transport in the halophilic archaeon *Halobacterium* sp. NRC-1: a major role for the twin-arginine translocation pathway? *Microbiology* **148**, 3335–3346.
5 Rose RW, Brüser T, Kissinger JC & Pohlschröder M (2002) Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway. *Mol Microbiol* **45**, 943–950.
6 Lanyi JK (1974) Salt-dependent properties of proteins from extremely halophilic bacteria. *Bacteriol Rev* **38**, 272–290.
7 Dilks K, Rose RW, Hartmann E & Pohlschröder M (2003) Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey. *J Bacteriol* **185**, 1478–1483.
8 Thomas JR & Bolhuis A (2006) The *tatC* gene cluster is essential for viability in halophilic archaea. *FEMS Microbiol Lett* **256**, 44–49.
9 Dilks K, Gimenez MI & Pohlschröder M (2005) Genetic and biochemical analysis of the twin-arginine translocation pathway in halophilic archaea. *J Bacteriol* **187**, 8104–8113.
10 Mould RM & Robinson C (1991) A proton gradient is required for the transport of two lumenal oxygen-evolving proteins across the thylakoid membrane. *J Biol Chem* **266**, 12189–12193.
11 Brau NA, Davis AW & Theg SM (2007) The chloroplast Tat pathway utilizes the transmembrane electric potential as an energy source. *Biophys J* **93**, 1993–1998.

12 Santini CL, Ize B, Chanal A, Muller M, Giordano G & Wu LF (1998) A novel sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J* **17**, 101–112.
13 Bageshwar UK & Musser SM (2007) Two electrical potential-dependent steps are required for transport by the *Escherichia coli* Tat machinery. *J Cell Biol* **179**, 87–99.
14 Hutcheon GW, Vasisht N & Bolhuis A (2005) Characterisation of a highly stable alpha-amylase from the halophilic archaeon *Haloarcula hispanica*. *Extremophiles* **9**, 487–495.
15 Muller V, Blaut M & Gottschalk G (1988) The transmembrane electrochemical gradient of Na^+ as driving force for methanol oxidation in *Methanosarcina barkeri*. *Eur J Biochem* **172**, 601–606.
16 Peddie CJ, Cook GM & Morgan HW (1999) Sodium-dependent glutamate uptake by an alkaliphilic, thermophilic *Bacillus* strain, TA2.A1. *J Bacteriol* **181**, 3172–3177.
17 Peddie CJ, Cook GM & Morgan HW (2000) Sucrose transport by the alkaliphilic, thermophilic *Bacillus* sp. strain TA2.A1 is dependent on a sodium gradient. *Extremophiles* **4**, 291–296.
18 Ring G, Londei P & Eichler J (2007) Protein biogenesis in Archaea: addressing translation initiation using an *in vitro* protein synthesis system for *Haloferax volcanii*. *FEMS Microbiol Lett* **270**, 34–41.
19 Bogsch EG, Sargent F, Stanley NR, Berks BC, Robinson C & Palmer T (1998) An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. *J Biol Chem* **273**, 18003–18006.
20 Sargent F, Stanley NR, Berks BC & Palmer T (1999) Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. *J Biol Chem* **274**, 36073–36082.
21 Jongbloed JD, Martin U, Antelmann H, Hecker M, Tjalsma H, Venema G, Bron S, van Dijl JM & Muller J (2000) TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway. *J Biol Chem* **275**, 41350–41357.
22 Oren A (1999) Bioenergetic aspects of halophilism. *Microbiol Mol Biol Rev* **63**, 334–348.
23 Hase CC (2003) Ion motive force dependence of protease secretion and phage transduction in *Vibrio cholerae* and *Pseudomonas aeruginosa*. *FEMS Microbiol Lett* **227**, 65–71.
24 Tokuda H, Kim YJ & Mizushima S (1990) *In vitro* protein translocation into inverted membrane vesicles prepared from *Vibrio alginolyticus* is stimulated by the electrochemical potential of Na^+ in the presence of *Escherichia coli* SecA. *FEBS Lett* **264**, 10–12.
25 Kamo N, Wakamatsu Y, Kohno K & Kobatake Y (1988) On the glutamate transport through cell envelope vesicles of *Halobacterium halobium*. *Biochem Biophys Res Commun* **152**, 1090–1096.
26 Lanyi JK, Renthal R & MacDonald RE (1976) Light-induced glutamate transport in *Halobacterium halobium* envelope vesicles. II. Evidence that the driving force is a light-dependent sodium gradient. *Biochemistry* **15**, 1603–1610.
27 Allers T, Ngo HP, Mevarech M & Lloyd RG (2004) Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. *Appl Environ Microbiol* **70**, 943–953.
28 Sambrook J & Russel D (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29 Cline SW, Lam WL, Charlebois RL, Schalkwyk LC & Doolittle WF (1989) Transformation methods for halophilic archaebacteria. *Can J Microbiol* **35**, 148–152.
30 Shi W, Tang XF, Huang Y, Gan F, Tang B & Shen P (2006) An extracellular halophilic protease SptA from a halophilic archaeon *Natrinema* sp. J7: gene cloning, expression and characterization. *Extremophiles* **10**, 599–606.
31 Steinert K, Wagner V, Kroth-Pancic PG & Bickel-Sandkotter S (1997) Characterization and subunit structure of the ATP synthase of the halophilic archaeon *Haloferax volcanii* and organization of the ATP synthase genes. *J Biol Chem* **272**, 6261–6269.
32 Ring G & Eichler J (2001) Characterization of inverted membrane vesicles from the halophilic archaeon *Haloferax volcanii*. *J Membr Biol* **183**, 195–204.

Bolhuis, A., Kwan, D., and Thomas, J.R. (2008) Halophilic adaptation in proteins. In: Protein adaptation in Extremophiles. Siddiqui, K.S. and Thomas, T. (Eds.) Nova Science Publishers, New York, Chapter 3, p71-104 [121]

References

1. Sargent, F., B.C. Berks, and T. Palmer, *Pathfinders and trailblazers: a prokaryotic targeting system for transport of folded proteins.* FEMS Microbiol Lett, 2006. **254**(2): p. 198-207.

2. Bolhuis, A., *The archaeal Sec-dependent protein translocation pathway.* Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 2004. **359**(1446): p. 919-927.

3. Pohlschroder, M., et al., *Protein translocation in the three domains of life: Variations on a theme.* Cell, 1997. **91**(5): p. 563-566.

4. Panzner, S., et al., *Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p.* Cell, 1995. **81**(4): p. 561-70.

5. Keenan, R.J., et al., *The signal recognition particle.* Annu Rev Biochem, 2001. **70**: p. 755-75.

6. Johnson, A.E. and M.A. van Waes, *The translocon: a dynamic gateway at the ER membrane.* Annu Rev Cell Dev Biol, 1999. **15**: p. 799-842.

7. Randall, L.L. and S.J. Hardy, *SecB, one small chaperone in the complex milieu of the cell.* Cell Mol Life Sci, 2002. **59**(10): p. 1617-23.

8. Hardy, S.J. and L.L. Randall, *Recognition of ligands by SecB, a molecular chaperone involved in bacterial protein export.* Philos Trans R Soc Lond B Biol Sci, 1993. **339**(1289): p. 343-52; discussion 352-4.

9. Manting, E.H. and A.J. Driessen, *Escherichia coli translocase: the unravelling of a molecular machine.* Mol Microbiol, 2000. **37**(2): p. 226-38.

10. Economou, A., et al., *SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF.* Cell, 1995. **83**(7): p. 1171-81.

11. Ulbrandt, N.D., J.A. Newitt, and H.D. Bernstein, *The E. coli signal recognition particle is required for the insertion of a subset of inner membrane proteins.* Cell, 1997. **88**(2): p. 187-96.

12. Seluanov, A. and E. Bibi, *FtsY, the prokaryotic signal recognition particle receptor homologue, is essential for biogenesis of membrane proteins.* J Biol Chem, 1997. **272**(4): p. 2053-5.

13. Netzer, W.J. and F.U. Hartl, *Recombination of protein domains facilitated by co-translational folding in eukaryotes.* Nature, 1997. **388**(6640): p. 343-9.

14. Brocchieri, L. and S. Karlin, *Protein length in eukaryotic and prokaryotic proteomes.* Nucleic Acids Res, 2005. **33**(10): p. 3390-400.

15. Ignatova, Z., et al., *From the test tube to the cell: exploring the folding and aggregation of a beta-clam protein.* Biopolymers, 2007. **88**(2): p. 157-63.

16. Eichler, J. and R. Moll, *The signal recognition particle of Archaea.* Trends Microbiol, 2001. **9**(3): p. 130-6.

17. Kinch, L.N., M.H. Saier, Jr., and N.V. Grishin, *Sec61beta--a component of the archaeal protein secretory system.* Trends Biochem Sci, 2002. **27**(4): p. 170-1.

18. Berks, B.C., T. Palmer, and F. Sargent, *Protein targeting by the bacterial twin-arginine translocation (Tat) pathway.* Current Opinion in Microbiology, 2005. **8**(2): p. 174-181.

19. DeLisa, M.P., D. Tullman, and G. Georgiou, *Folding quality control in the export of proteins by the bacterial twin-arginine translocation pathway.* Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(10): p. 6115-6120.

20. Hynds, P.J., D. Robinson, and C. Robinson, *The sec-independent twin-arginine translocation system can transport both tightly folded and malfolded proteins across the thylakoid membrane.* J Biol Chem, 1998. **273**(52): p. 34868-74.

21. Bolhuis, A., *Protein transport in the halophilic archaeon ARTICLE Halobacterium sp NRC-1: a major role for the twin-arginine translocation pathway?* Microbiology-Sgm, 2002. **148**: p. 3335-3346.

22. Berks, B.C., *A common export pathway for proteins binding complex redox cofactors?* Mol Microbiol, 1996. **22**(3): p. 393-404.

23. Robinson, C. and A. Bolhuis, *Protein targeting by the twin-arginine translocation pathway.* Nature Reviews Molecular Cell Biology, 2001. **2**(5): p. 350-356.

24. Ize, B., et al., *Role of the Escherichia coli Tat pathway in outer membrane integrity.* Mol Microbiol, 2003. **48**(5): p. 1183-93.

25. Oates, J., et al., *Consensus structural features of purified bacterial TatABC complexes.* Journal of Molecular Biology, 2003. **330**(2): p. 277-286.

26. Jack, R.L., et al., *Constitutive expression of Escherichia coli tat genes indicates an important role for the twin-arginine translocase during aerobic and anaerobic growth.* Journal of Bacteriology, 2001. **183**(5): p. 1801-1804.

27. Stanley, N.R., et al., *Escherichia coli strains blocked in Tat-dependent protein export exhibit pleiotropic defects in the cell envelope.* Journal of Bacteriology, 2001. **183**(1): p. 139-144.

28. Berg, B.L., et al., *Nitrate-inducible formate dehydrogenase in Escherichia coli K-12. I. Nucleotide sequence of the fdnGHI operon and evidence that opal (UGA) encodes selenocysteine.* J Biol Chem, 1991. **266**(33): p. 22380-5.

29. Weiner, J.H., et al., *A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins.* Cell, 1998. **93**(1): p. 93-101.

30. Hussain, H., et al., *A seven-gene operon essential for formate-dependent nitrite reduction to ammonia by enteric bacteria.* Mol Microbiol, 1994. **12**(1): p. 153-63.

31. von Heijne, G., *The signal peptide.* J Membr Biol, 1990. **115**: p. 195-201.

32. Cristobal, S., et al., *Competition between Sec- and TAT-dependent protein translocation in Escherichia coli.* Embo Journal, 1999. **18**(11): p. 2982-2990.

33. Ize, B., F. Gerard, and L.F. Wu, *In vivo assessment of the Tat signal peptide specificity in Escherichia coli.* Archives of Microbiology, 2002. **178**(6): p. 548-553.

34. Bogsch, E., S. Brink, and C. Robinson, *Pathway specificity for a delta pH-dependent precursor thylakoid lumen protein is governed by a 'Sec-avoidance' motif in the transfer peptide and a 'Sec-incompatible' mature protein.* Embo J, 1997. **16**(13): p. 3851-9.

35. Chaddock, A.M., et al., *A new type of signal peptide: central role of a twin-arginine motif in transfer signals for the delta pH-dependent thylakoidal protein translocase.* Embo J, 1995. **14**(12): p. 2715-22.

36. Halbig, D., et al., *The efficient export of NADP-containing glucose-fructose oxidoreductase to the periplasm of Zymomonas mobilis depends both on an intact twin-arginine motif in the signal peptide and on the generation of a structural export signal induced by cofactor binding.* Eur J Biochem, 1999. **263**(2): p. 543-51.

37. Ize, B., et al., *In vivo dissection of the tat translocation pathway in Escherichia coli.* Journal of Molecular Biology, 2002. **317**(3): p. 327-335.

38. Alami, M., et al., *Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli.* Molecular Cell, 2003. **12**(4): p. 937-946.

39. de Leeuw, E., et al., *Oligomeric properties and signal peptide binding by Escherichia coli Tat protein transport complexes.* Journal of Molecular Biology, 2002. **322**(5): p. 1135-1146.

40. Jormakka, M., et al., *Molecular basis of proton motive force generation: structure of formate dehydrogenase-N.* Science, 2002. **295**(5561): p. 1863-8.

41. Stanley, N.R., et al., *Behaviour of topological marker proteins targeted to the Tat protein transport pathway.* Mol Microbiol, 2002. **43**(4): p. 1005-21.

42. Jack, R.L., et al., *Coordinating assembly and export of complex bacterial proteins.* Embo J, 2004. **23**(20): p. 3962-72.

43. Rodrigue, A., et al., *Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway.* J Biol Chem, 1999. **274**(19): p. 13223-8.

44. Barrett, C.M.L., et al., *Quantitative export of a reporter protein, GFP, by the twin-arginine translocation pathway in Escherichia coli.* Biochemical and Biophysical Research Communications, 2003. **304**(2): p. 279-284.

45. Yahr, T.L. and W.T. Wickner, *Functional reconstitution of bacterial Tat translocation in vitro.* Embo Journal, 2001. **20**(10): p. 2472-2479.

46. Wexler, M., et al., *TatD is a cytoplasmic protein with DNase activity. No requirement for TatD family proteins in sec-independent protein export.* J Biol Chem, 2000. **275**(22): p. 16717-22.

47. Matos, C.F., A. Di Cola, and C. Robinson, *TatD is a central component of a Tat translocon-initiated quality control system for exported FeS proteins in Escherichia coli.* EMBO Rep, 2009. **10**(5): p. 474-9.

48. Berks, B.C., F. Sargent, and T. Palmer, *The Tat protein export pathway.* Molecular Microbiology, 2000. **35**(2): p. 260-274.

49. Lee, P.A., et al., *Truncation analysis of TatA and TatB defines the minimal functional units required for protein translocation.* Journal of Bacteriology, 2002. **184**(21): p. 5871-5879.

50. Porcelli, I., et al., *Characterization and membrane assembly of the TatA component of the Escherichia coli twin-arginine protein transport system.* Biochemistry, 2002. **41**(46): p. 13690-13697.

51. Sargent, F., et al., *Overlapping functions of components of a bacterial Sec-independent protein export pathway.* Embo Journal, 1998. **17**(13): p. 3640-3650.

52. Palmer, T., F. Sargent, and B.C. Berks, *Export of complex cofactor-containing proteins by the bacterial Tat pathway.* Trends in Microbiology, 2005. **13**(4): p. 175-180.

53. Sargent, F., et al., *Sec-independent protein translocation in Escherichia coli - A distinct and pivotal role for the TatB protein.* Journal of Biological Chemistry, 1999. **274**(51): p. 36073-36082.

54. Bolhuis, A., et al., *TatB and TatC form a functional and structural unit of the twin-arginine translocase from Escherichia coli.* Journal of Biological Chemistry, 2001. **276**(23): p. 20213-20219.

55. Sargent, F., et al., *Purified components of the Escherichia coli Tat protein transport system form a double-layered ring structure.* European Journal of Biochemistry, 2001. **268**(12): p. 3361-3367.

56. De Leeuw, E., et al., *Membrane interactions and self-association of the TatA and TatB components of the twin-arginine translocation pathway.* Febs Letters, 2001. **506**(2): p. 143-148.

57. Muller, M. and R.B. Klosgen, *The Tat pathway in bacteria and chloroplasts (review).* Mol Membr Biol, 2005. **22**(1-2): p. 113-21.

58. Alami, M., et al., *Separate analysis of twin-arginine translocation (Tat)-specific membrane binding and translocation in Escherichia coli.* J Biol Chem, 2002. **277**(23): p. 20499-503.

59. Cruz, J.A., et al., *Contribution of electric field (Delta psi) to steady-state transthylakoid proton motive force (pmf) in vitro and in vivo. control of pmf parsing into Delta psi and Delta pH by ionic strength.* Biochemistry, 2001. **40**(5): p. 1226-37.

60. Pugsley, A.P., *The complete general secretory pathway in gram-negative bacteria.* Microbiol Rev, 1993. **57**(1): p. 50-108.

61. Pradel, N., et al., *Contribution of the twin arginine translocation system to the virulence of enterohemorrhagic Escherichia coli O157 : H7.* Infection and Immunity, 2003. **71**(9): p. 4908-4916.

62. Jakob, M., et al., *Tat subunit stoichiometry in Arabidopsis thaliana challenges the proposed function of TatA as the translocation pore.* Biochim Biophys Acta, 2008.

63. Gouffi, K., et al., *Dual topology of the Escherichia coli TatA protein.* Journal of Biological Chemistry, 2004. **279**(12): p. 11608-11615.

64. Natale, P., T. Bruser, and A.J. Driessen, *Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms.* Biochim Biophys Acta, 2008. **1778**(9): p. 1735-56.

65. Alder, N.N. and S.M. Theg, *Energetics of protein transport across biological membranes. a study of the thylakoid DeltapH-dependent/cpTat pathway.* Cell, 2003. **112**(2): p. 231-42.

66. Gohlke, U., et al., *The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(30): p. 10482-10486.

67. Schaerlaekens, K., et al., *Twin-arginine translocation pathway in Streptomyces lividans.* J Bacteriol, 2001. **183**(23): p. 6727-32.

68. Blaudeck, N., et al., *Isolation and characterization of bifunctional Escherichia coli TatA mutant proteins that allow efficient Tat-dependent protein translocation in the absence of TatB.* Journal of Biological Chemistry, 2005. **280**(5): p. 3426-3432.

69. Barnett, J.P., et al., *A minimal tat system from a gram-positive organism - A bifunctional TatA subunit participates in discrete TatAC and TatA complexes.* Journal of Biological Chemistry, 2008. **283**(5): p. 2534-2542.

70. Pop, O., et al., *The twin-arginine signal peptide of PhoD and the TatA(d)/C-d proteins of Bacillus subtilis form an autonomous tat translocation system.* Journal of Biological Chemistry, 2002. **277**(5): p. 3268-3273.

71. Eder, S., et al., *A Bacillus subtilis secreted phosphodiesterase/alkaline phosphatase is the product of a Pho regulon gene, phoD.* Microbiology, 1996. **142 (Pt 8)**: p. 2041-7.

72. Jongbloed, J.D.H., et al., *TatC is a specificity determinant for protein secretion via the twin-arginine translocation pathway.* Journal of Biological Chemistry, 2000. **275**(52): p. 41350-41357.

73. Jongbloed, J.D., et al., *Two minimal Tat translocases in Bacillus.* Mol Microbiol, 2004. **54**(5): p. 1319-25.

74. Barnett, J.P., et al., *The twin-arginine translocation (Tat) systems from Bacillus subtilis display a conserved mode of complex organization and similar substrate recognition requirements.* Febs J, 2009. **276**(1): p. 232-43.

75. Widdick, D.A., et al., *A facile reporter system for the experimental identification of twin-arginine translocation (Tat) signal peptides from all kingdoms of life.* Journal of Molecular Biology, 2008. **375**(3): p. 595-603.

76. De Keersmaeker, S., et al., *Structural organization of the twin-arginine translocation system in Streptomyces lividans.* FEBS Lett, 2005. **579**(3): p. 797-802.

77. Westermann, M., et al., *The TatA(d) component of the Bacillus subtilis twin-arginine protein transport system forms homo-multimeric complexes in its cytosolic and membrane embedded localisation.* Biochim Biophys Acta, 2006.

78. Pop, O.I., et al., *Sequence-specific binding of prePhoD to soluble TatAd indicates protein-mediated targeting of the Tat export in Bacillus subtilis.* J Biol Chem, 2003. **278**(40): p. 38428-36.

79. Chanal, A., C.L. Santini, and L.F. Wu, *Specific inhibition of the translocation of a subset of Escherichia coli TAT substrates by the TorA signal peptide.* J Mol Biol, 2003. **327**(3): p. 563-70.

80. Chatton, F., *Titres et Travoux Scientifiques.* de Edouard Chatton, 1938: p. 1906-1937.

81. Woese, C.R. and G.E. Fox, *Phylogenetic structure of the prokaryotic domain: the primary kingdoms.* Proc Natl Acad Sci U S A, 1977. **74**(11): p. 5088-90.

82. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4576-9.

83. Stetter, K.O., *History of discovery of the first hyperthermophiles.* Extremophiles, 2006. **10**(5): p. 357-62.

84. DeLong, E.F., et al., *High abundance of Archaea in Antarctic marine picoplankton.* Nature, 1994. **371**(6499): p. 695-7.

85. Huber, H., et al., *A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont.* Nature, 2002. **417**(6884): p. 63-7.

86. Baldauf, S.L., J.D. Palmer, and W.F. Doolittle, *The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny.* Proc Natl Acad Sci U S A, 1996. **93**(15): p. 7749-54.

87. Yang, Z. and D. Roberts, *On the use of nucleic acid sequences to infer early branchings in the tree of life.* Mol Biol Evol, 1995. **12**(3): p. 451-8.

88. Rivera, M.C. and J.A. Lake, *Evidence that eukaryotes and eocyte prokaryotes are immediate relatives.* Science, 1992. **257**(5066): p. 74-6.

89. Cox, C.J., et al., *The archaebacterial origin of eukaryotes.* Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20356-61.

90. Karner, M.B., E.F. DeLong, and D.M. Karl, *Archaeal dominance in the mesopelagic zone of the Pacific Ocean.* Nature, 2001. **409**(6819): p. 507-10.

91. De Rosa M and G. A, *The lipids of archaebacteria.* Prog. Lipid Res, 1888. **27**: p. 153-175.

92. Nishihara, M., et al., *Archaea contain a novel diether phosphoglycolipid with a polar head group identical to the conserved core of eucaryal glycosyl phosphatidylinositol.* J Biol Chem, 1992. **267**(18): p. 12432-5.

93. Yen, M.R., et al., *Sequence and phylogenetic analyses of the twin-arginine targeting (Tat) protein export system.* Arch Microbiol, 2002. **177**(6): p. 441-50.

94. Pohlschroder, M., et al., *Translocation of proteins across archaeal cytoplasmic membranes.* Fems Microbiology Reviews, 2004. **28**(1): p. 3-24.

95. van de Vossenberg, J.L.C.M., A.J.M. Driessen, and W.N. Konings, *The essence of being extremophilic: the role of the unique archaeal membrane lipids.* Extremophiles, 1998. **2**(3): p. 163-170.

96. Brock, T.D., et al., *Sulfolobus: a new genus of sulfur-oxidizing bacteria living at low pH and high temperature.* Arch Mikrobiol, 1972. **84**(1): p. 54-68.

97. Chen, L., et al., *The genome of Sulfolobus acidocaldarius, a model organism of the Crenarchaeota.* J Bacteriol, 2005. **187**(14): p. 4992-9.

98. Kawarabayasi, Y., et al., *Complete genome sequence of an aerobic thermoacidophilic crenarchaeon, Sulfolobus tokodaii strain7.* DNA Res, 2001. **8**(4): p. 123-40.

99. She, Q., et al., *The complete genome of the crenarchaeon Sulfolobus solfataricus P2.* Proc Natl Acad Sci U S A, 2001. **98**(14): p. 7835-40.

100. Auernik, K.S., et al., *The genome sequence of the metal-mobilizing, extremely thermoacidophilic archaeon Metallosphaera sedula provides insights into bioleaching-associated metabolism.* Appl Environ Microbiol, 2008. **74**(3): p. 682-92.

101. De Rosa, M., et al., *Glucose metabolism in the extreme thermoacidophilic archaebacterium Sulfolobus solfataricus.* Biochem J, 1984. **224**(2): p. 407-14.

102. Grogan, D.W., *Phenotypic characterization of the archaebacterial genus Sulfolobus: comparison of five wild-type strains.* J Bacteriol, 1989. **171**(12): p. 6710-9.

103. Suzuki, T., et al., *Sulfolobus tokodaii sp. nov. (f. Sulfolobus sp. strain 7), a new member of the genus Sulfolobus isolated from Beppu Hot Springs, Japan.* Extremophiles, 2002. **6**(1): p. 39-44.

104. Iwasaki, T., et al., *Resolution of the aerobic respiratory system of the thermoacidophilic archaeon, Sulfolobus sp. strain 7. II. Characterization of the archaeal terminal oxidase subcomplexes and implication for the intramolecular electron transfer.* J Biol Chem, 1995. **270**(52): p. 30893-901.

105. Konings, W.N., et al., *The cell membrane plays a crucial role in survival of bacteria and archaea in extreme environments.* Antonie Van Leeuwenhoek International Journal of General and Molecular Microbiology, 2002. **81**(1-4): p. 61-72.

106. Nicolaus, B., et al., *Calditol tetraether lipids of the archaebacterium Sulfolobus solfataricus. Biosynthetic studies.* Biochem J, 1990. **266**(3): p. 785-91.

107. Bagatolli, L., et al., *Two-photon fluorescence microscopy studies of bipolar tetraether giant liposomes from thermoacidophilic archaebacteria Sulfolobus acidocaldarius.* Biophys J, 2000. **79**(1): p. 416-25.

108. Albers, S.V. and A.J.M. Driessen, *Signal peptides of secreted proteins of the archaeon Sulfolobus solfataricus: a genomic survey.* Archives of Microbiology, 2002. **177**(3): p. 209-216.

109. Hutcheon, G.W. and A. Bolhuis, *The archaeal twin-arginine translocation pathway.* Biochemical Society Transactions, 2003. **31**: p. 686-689.

110. Gurung, B., et al., *The iron-sulfur cluster of the Rieske iron-sulfur protein functions as a proton-exiting gate in the cytochrome bc(1) complex.* J Biol Chem, 2005. **280**(26): p. 24895-902.

111. Hilton, J.C. and K.V. Rajagopalan, *Identification of the molybdenum cofactor of dimethyl sulfoxide reductase from Rhodobacter sphaeroides f. sp. denitrificans as bis(molybdopterin guanine dinucleotide)molybdenum.* Arch Biochem Biophys, 1996. **325**(1): p. 139-43.

112. Schmidt, C.L., *Rieske iron-sulfur proteins from extremophilic organisms.* J Bioenerg Biomembr, 2004. **36**(1): p. 107-13.

113. Bohuis, A., G. Hutcheon, and N. Vasisht, *Characterisation of alpha-amylase from the halophillic archaeon Haloarcula hispania.* Springer, 2005.

114. Lanyi, J.K., *Salt-dependent properties of proteins from extremely halophilic bacteria.* Bacteriol Rev, 1974. **38**(3): p. 272-90.

115. Oren, A., *Microbial life at high salt concentrations: phylogenetic and metabolic diversity.* Saline Systems, 2008. **4**: p. 2.

116. Madern, D., C. Ebel, and G. Zaccai, *Halophilic adaptation of enzymes.* Extremophiles, 2000. **4**(2): p. 91-8.

117. Dilks, K., et al., *Prokaryotic utilization of the twin-arginine translocation pathway: a genomic survey.* J Bacteriol, 2003. **185**(4): p. 1478-83.

118. Rose, R.W., et al., *Adaptation of protein secretion to extremely high-salt conditions by extensive use of the twin-arginine translocation pathway.* Mol Microbiol, 2002. **45**(4): p. 943-50.

119. Thomas, J.R. and A. Bolhuis, *The tatC gene cluster is essential for viability in halophilic archaea.* FEMS Microbiol Lett, 2006. **256**(1): p. 44-9.

120. Kennedy, S.P., et al., *Understanding the adaptation of Halobacterium species NRC-1 to its extreme environment through computational analysis of its genome sequence.* Genome Res, 2001. **11**(10): p. 1641-50.

121. Bolhuis, A., D. Kwan, and J.R. Thomas, *Halophilic adaptation in proteins.* . In: Protein adaptation in Extremophiles. Siddiqui, K.S. and Thomas, T. (Eds.), Nova Science Publishers, New York, 2008: p. 71-104.

122. Mevarech, M., F. Frolow, and L.M. Gloss, *Halophilic enzymes: proteins with a grain of salt.* Biophys Chem, 2000. **86**(2-3): p. 155-64.

123. Franzetti, B., et al., *Characterization of a novel complex from halophilic archaebacteria, which displays chaperone-like activities in vitro.* J Biol Chem, 2001. **276**(32): p. 29906-14.

124. Ng, W.V., et al., *Genome sequence of Halobacterium species NRC-1.* Proc Natl Acad Sci U S A, 2000. **97**(22): p. 12176-81.

125. Dilks, K., M.I. Gimenez, and M. Pohlschroder, *Genetic and biochemical analysis of the twin-arginine translocation pathway in halophilic archaea.* J Bacteriol, 2005. **187**(23): p. 8104-13.

126. Juez, G., et al., *Haloarcula hispanica spec. nov. and Haloferax gibbonsii spec. nov., two new species of extremely halophillic archaeabacteria.* Syst Appl Microbiol, 1986. **8**: p. 75-79.

127. Sowers, K.R. and H.J. Schreier, *Gene transfer systems for the Archaea.* Trends Microbiol, 1999. **7**(5): p. 212-9.

128. Peck, R.F., S. Dassarma, and M.P. Krebs, *Homologous gene knockout in the archaeon Halobacterium salinarum with ura3 as a counterselectable marker.* Mol Microbiol, 2000. **35**(3): p. 667-76.

129. Guzman, L.M., et al., *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* J Bacteriol, 1995. **177**(14): p. 4121-30.

130. Yang, Y., Y.P. Huang, and P. Shen, *The 492-bp RM07 DNA fragment from the halophilic Archaea confers promoter activity in all three domains of life.* Curr Microbiol, 2003. **47**(5): p. 388-94.

131. Albers, S.V., et al., *Production of recombinant and tagged proteins in the Hyperthermophillic Archaeon Sulfolobus solfataricus.* Applied and environmental microbiology, 2006. **72**(1): p. 1-8.

132. Albers, S.V. and A.J. Driessen, *Conditions for gen disruption by homologous recombination of exogenous DNA into the Sulfolobus solfataricus genome.* Archaea, 2007. **2**: p. 145-149.

133. Webb, B.L., M.M. Cox, and R.B. Inman, *An interaction between the Escherichia coli RecF and RecR proteins dependent on ATP and double-stranded DNA.* J Biol Chem, 1995. **270**(52): p. 31397-404.

134. Casadaban, M.J., *Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu.* J Mol Biol, 1976. **104**(3): p. 541-55.

135. Yanisch-Perron, C., J. Vieira, and J. Messing, *Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors.* Gene, 1985. **33**(1): p. 103-19.

136. Dower, W.J., J.F. Miller, and C.W. Ragsdale, *High efficiency transformation of E. coli by high voltage electroporation.* Nucleic Acids Res, 1988. **16**(13): p. 6127-45.

137. Kowit, J.D. and A.L. Goldberg, *Intermediate steps in the degradation of a specific abnormal protein in Escherichia coli.* J Biol Chem, 1977. **252**(23): p. 8350-7.

138. Palmer, B.R. and M.G. Marinus, *The dam and dcm strains of Escherichia coli--a review.* Gene, 1994. **143**(1): p. 1-12.

139. Trauger, J.W. and C.T. Walsh, *Heterologous expression in Escherichia coli of the first module of the nonribosomal peptide synthetase for chloroeremomycin, a vancomycin-type glycopeptide antibiotic.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3112-7.

140. Allers, T., et al., *Development of additional selectable markers for the halophilic archaeon Haloferax volcanii based on the leuB and trpA genes.* Appl Environ Microbiol, 2004. **70**(2): p. 943-53.

141. Albers, S.V., et al., *Production of recombinant and tagged proteins in the hyperthermophilic archaeon Sulfolobus solfataricus.* Appl Environ Microbiol, 2006. **72**(1): p. 102-11.

142. Worthington, P., et al., *Targeted disruption of the alpha-amylase gene in the hyperthermophilic archaeon Sulfolobus solfataficus.* Journal of Bacteriology, 2003. **185**(2): p. 482-488.

143. Jones, B.E., et al., *Microbial diversity of soda lakes.* Extremophiles, 1998. **2**(3): p. 191-200.

144. Sambrook, J. and D. Russell, *Molecular Cloning: A Laboratory Manual (3-Volume Set)*. 1989.

145. Rolfsmeier, M. and P. Blum, *Purification and characterization of a maltase from the extremely thermophilic crenarchaeote Sulfolobus solfataricus.* J Bacteriol, 1995. **177**(2): p. 482-5.

146. Yeats, S., P. McWilliam, and W. Zillig, *A plasmid in the archaebacterium Sulfolobus acidocaldarius.* Embo J, 1982. **1**(9): p. 1035-1038.

147. Schelert, J., et al., *Occurrence and characterization of mercury resistance in the hyperthermophilic archaeon Sulfolobus solfataricus by use of gene disruption.* J Bacteriol, 2004. **186**(2): p. 427-37.

148. Cline, S.W., et al., *Transformation methods for halophilic archaebacteria.* Can J Microbiol, 1989. **35**(1): p. 148-52.

149. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.

150. Hutcheon, G., A. Bolhuis, and N. Vasisht, *Characterisation of alpha-amylase from the halophillic archaeon Haloarcula hispania.* Springer, 2005.

151. Randall, L.L. and S.J. Hardy, *Correlation of competence for export with lack of tertiary structure of the mature species: a study in vivo of maltose-binding protein in E. coli.* Cell, 1986. **46**(6): p. 921-8.

152. Kwan, D.C., J.R. Thomas, and A. Bolhuis, *Bioenergetic requirements of a Tat-dependent substrate in the halophilic archaeon Haloarcula hispanica.* Febs J, 2008. **275**(24): p. 6159-67.

153. Sensen, C.W., et al., *Completing the sequence of the Sulfolobus solfataricus P2 genome.* Extremophiles, 1998. **2**(3): p. 305-12.

154. El Karoui, M., et al., *Gene replacement with linear DNA in electroporated wild-type Escherichia coli.* Nucleic Acids Research, 1999. **27**(5): p. 1296-1299.

155. Allers, T. and M. Mevarech, *Archaeal genetics - The third way.* Nature Reviews Genetics, 2005. **6**(1): p. 58-73.

156. Rother, M. and W.W. Metcalf, *Genetic technologies for Archaea.* Current Opinion in Microbiology, 2005. **8**(6): p. 745-751.

157. Seonghun, K. and S.B. Lee, *Rare codon clusters at 5'-end influence heterologous expression of archaeal gene in Escherichia coli.* Protein expression and purification 2006. **50**: p. 49-57.

158. Kim, S. and S.B. Lee, *Rare codon clusters at 5'-end influence heterologous expression of archaeal gene in Escherichia coli.* Protein Expr Purif, 2006. **50**(1): p. 49-57.

159. Prisco, A., et al., *A gene encoding a putative membrane protein homologous to the major facilitator superfamily of transporters maps upstream of the beta-glycosidase gene in the archaeon Sulfolobus solfataricus.* J Bacteriol, 1995. **177**(6): p. 1614-9.

160. Van Voorst, F. and B. De Kruijff, *Role of lipids in the translocation of proteins across membranes.* Biochem J, 2000. **347 Pt 3**: p. 601-12.

161. Moracci, M., et al., *Expression of the thermostable beta-galactosidase gene from the archaebacterium Sulfolobus solfataricus in Saccharomyces cerevisiae and characterization of a new inducible promoter for heterologous expression.* J Bacteriol, 1992. **174**(3): p. 873-82.

162. Hjelmeland, L.M. and A. Chrambach, *Solubilization of functional membrane proteins.* Methods Enzymol, 1984. **104**: p. 305-18.

163. Hicks, M.G., et al., *Formation of functional Tat translocases from heterologous components.* BMC Microbiol, 2006. **6**: p. 64.

164. Xiong, Y., et al., *Expression level of heterologous tat genes is crucial for in vivo reconstitution of a functional Tat translocase in Escherichia coli.* Biochimie, 2007. **89**(5): p. 676-85.

165. Saint-Joanis, B., et al., *Inactivation of Rv2525c, a substrate of the twin arginine translocation (Tat) system of Mycobacterium tuberculosis, increases beta-lactam susceptibility and virulence.* J Bacteriol, 2006. **188**(18): p. 6669-79.

166. Rothery, R.A. and J.H. Weiner, *Alteration of the iron-sulfur cluster composition of Escherichia coli dimethyl sulfoxide reductase by site-directed mutagenesis.* Biochemistry, 1991. **30**(34): p. 8296-305.

167. Bilous, P.T. and J.H. Weiner, *Molecular cloning and expression of the Escherichia coli dimethyl sulfoxide reductase operon.* J Bacteriol, 1988. **170**(4): p. 1511-8.

168. Barrett, E.L. and H.S. Kwan, *Bacterial reduction of trimethylamine oxide.* Annu Rev Microbiol, 1985. **39**: p. 131-49.

169. Santini, C.L., et al., *A novel sec-independent periplasmic protein translocation pathway in Escherichia coli.* Embo J, 1998. **17**(1): p. 101-12.

170. Bolhuis, A., *Protein transport in the halophilic archaeon Halobacterium sp. NRC-1: a major role for the twin-arginine translocation pathway?* Microbiology, 2002. **148**(Pt 11): p. 3335-46.

171. Rose, R.W. and M. Pohlschroder, *In vivo analysis of an essential archaeal signal recognition particle in its native host.* Journal of Bacteriology, 2002. **184**(12): p. 3260-3267.

172. Shi, W., et al., *An extracellular halophilic protease SptA from a halophilic archaeon Natrinema sp. J7: gene cloning, expression and characterization.* Extremophiles, 2006. **10**(6): p. 599-606.

173. Henry, R., et al., *Targeting determinants and proposed evolutionary basis for the Sec and the Delta pH protein transport systems in chloroplast thylakoid membranes.* J Cell Biol, 1997. **136**(4): p. 823-32.

174. Stanley, N.R., T. Palmer, and B.C. Berks, *The twin arginine consensus motif of Tat signal peptides is involved in Sec-independent protein targeting in Escherichia coli.* Journal of Biological Chemistry, 2000. **275**(16): p. 11591-11596.

175. Ring, G. and J. Eichler, *Characterization of inverted membrane vesicles from the halophilic archaeon Haloferax volcanii.* Journal of Membrane Biology, 2001. **183**(3): p. 195-204.

176. Hutcheon, G.W., *Protein secretion in halophilic archaea.* Thesis (M. Phil.), 2008. **University of Warwick**.

177. Richter, S., et al., *Functional Tat transport of unstructured, small, hydrophilic proteins.* J Biol Chem, 2007. **282**(46): p. 33257-64.

178. Nagashima, T., et al., *Site-directed mutagenesis of catalytic active-site residues of Taka-amylase A.* Biosci Biotechnol Biochem, 1992. **56**(2): p. 207-10.

179. Rodriguez-Sanoja, R., N. Oviedo, and S. Sanchez, *Microbial starch-binding domain.* Curr Opin Microbiol, 2005. **8**(3): p. 260-7.

180. Dreusch, A., et al., *Lack of copper insertion into unprocessed cytoplasmic nitrous oxide reductase generated by an R20D substitution in the arginine consensus motif of the signal peptide.* Biochim Biophys Acta, 1997. **1319**(2-3): p. 311-8.

181. Takase, K., *Site-directed mutagenesis reveals critical importance of the catalytic site in the binding of alpha-amylase by wheat proteinaceous inhibitor.* Biochemistry, 1994. **33**(25): p. 7925-30.

182. Bageshwar, U.K. and S.M. Musser, *Two electrical potential-dependent steps are required for transport by the Escherichia coli Tat machinery.* J Cell Biol, 2007. **179**(1): p. 87-99.

183. Keegstra, K. and K. Cline, *Protein import and routing systems of chloroplasts.* Plant Cell, 1999. **11**(4): p. 557-70.

184. Braun, N.A., A.W. Davis, and S.M. Theg, *The chloroplast Tat pathway utilizes the transmembrane electric potential as an energy source.* Biophys J, 2007. **93**(6): p. 1993-8.

185. Peddie, C.J., G.M. Cook, and H.W. Morgan, *Sucrose transport by the alkaliphilic, thermophilic Bacillus sp. strain TA2.A1 is dependent on a sodium gradient.* Extremophiles, 2000. **4**(5): p. 291-6.

186. Ahmed, S. and I.R. Booth, *The use of valinomycin, nigericin and trichlorocarbanilide in control of the protonmotive force in Escherichia coli cells.* Biochem J, 1983. **212**(1): p. 105-12.

187. Horikoshi, K., *Alkaliphiles: some applications of their products for biotechnology.* Microbiol Mol Biol Rev, 1999. **63**(4): p. 735-50, table of contents.

188. Booth, I.R., *Regulation of cytoplasmic pH in bacteria.* Microbiol Rev, 1985. **49**(4): p. 359-78.

189. van de Vossenberg, J.L., et al., *Lipid membranes from halophilic and alkali-halophilic Archaea have a low H+ and Na+ permeability at high salt concentration.* Extremophiles, 1999. **3**(4): p. 253-7.

190. Tindall, B.J., H.N.M. Ross, and W.D. Grant, *Natronobacterium Gen-Nov and Natronococcus Gen-Nov, 2 New Genera of Haloalkaliphilic Archaebacteria.* Systematic and Applied Microbiology, 1984. **5**(1): p. 41-57.

191. Soliman, G.S.H. and H.G. Trüper, *Halobacterium pharaonis sp. nov., a new extremely haloalkaliphilic archaebacterium with low magnesium requirement.* Zbl. Bakt. Hyg. I Abt. Orig., 1982. **3**: p. 318–329.

192. Cao, Y., et al., *Characterization of alcohol dehydrogenase from the haloalkaliphilic archaeon Natronomonas pharaonis.* Extremophiles, 2008. **12**(3): p. 471-6.

193. Muller, V., M. Blaut, and G. Gottschalk, *The transmembrane electrochemical gradient of Na+ as driving force for methanol oxidation in Methanosarcina barkeri.* Eur J Biochem, 1988. **172**(3): p. 601-6.

194. Peddie, C.J., G.M. Cook, and H.W. Morgan, *Sodium-dependent glutamate uptake by an alkaliphilic, thermophilic Bacillus strain, TA2.A1.* J Bacteriol, 1999. **181**(10): p. 3172-7.

195. Kamo, N., et al., *On the glutamate transport through cell envelope vesicles of Halobacterium halobium.* Biochem Biophys Res Commun, 1988. **152**(3): p. 1090-6.

196. Hase, C.C., *Ion motive force dependence of protease secretion and phage transduction in Vibrio cholerae and Pseudomonas aeruginosa.* FEMS Microbiol Lett, 2003. **227**(1): p. 65-71.

197. Tokuda, H., Y.J. Kim, and S. Mizushima, *In vitro protein translocation into inverted membrane vesicles prepared from Vibrio alginolyticus is stimulated by the electrochemical potential of Na+ in the presence of Escherichia coli SecA.* FEBS Lett, 1990. **264**(1): p. 10-2.

198. Krulwich, T.A. and A.A. Guffanti, *The Na+ cycle of extreme alkalophiles: a secondary Na+/H+ antiporter and Na+/solute symporters.* J Bioenerg Biomembr, 1989. **21**(6): p. 663-77.

199. Kobayashi, T., et al., *Haloalkaliphilic maltotriose-forming alpha-amylase from the archaebacterium Natronococcus sp. strain Ah-36.* J Bacteriol, 1992. **174**(11): p. 3439-44.

200. Good, W.A. and P.A. Hartman, *Properties of the amylase from Halobacterium halobium.* J Bacteriol, 1970. **104**(1): p. 601-3.

201. Perez-Pomares, F., et al., *Alpha-amylase activity from the halophilic archaeon Haloferax mediterranei.* Extremophiles, 2003. **7**(4): p. 299-306.

202. Fukushima, T., et al., *Organic solvent tolerance of halophilic alpha-amylase from a Haloarchaeon, Haloarcula sp. strain S-1.* Extremophiles, 2005. **9**(1): p. 85-9.

203. Kobayashi, T., et al., *Cloning, expression, and nucleotide sequence of the alpha-amylase gene from the haloalkaliphilic archaeon Natronococcus sp. strain Ah-36.* J Bacteriol, 1994. **176**(16): p. 5131-4.

204. Cleland, W.W., *Dithiothreitol, a New Protective Reagent for Sh Groups.* Biochemistry, 1964. **3**: p. 480-2.

205. Mijts, B.N. and B.K. Patel, *Cloning, sequencing and expression of an alpha-amylase gene, amyA, from the thermophilic halophile Halothermothrix orenii and purification and biochemical characterization of the recombinant enzyme.* Microbiology, 2002. **148**(Pt 8): p. 2343-9.

206. Benlloch, S., et al., *Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern.* Environ Microbiol, 2002. **4**(6): p. 349-60.

207. Aitken, D.M. and A.D. Brown, *Citrate and glyoxylate cycles in the halophil, Halobacterium salinarium.* Biochim Biophys Acta, 1969. **177**(2): p. 351-4.

208. Bayley, S.T. and D.J. Kushner, *The Ribosomes of the Extremely Halophilic Bacterium, Halobacterium Cutirubrum.* J Mol Biol, 1964. **9**: p. 654-69.

209. Bonete, M.J., et al., *Glucose dehydrogenase from the halophilic Archaeon Haloferax mediterranei: enzyme purification, characterisation and N-terminal sequence.* FEBS Lett, 1996. **383**(3): p. 227-9.

210. Martinez-Espinosa, R.M., et al., *An octameric prokaryotic glutamine synthetase from the haloarchaeon Haloferax mediterranei.* FEMS Microbiol Lett, 2006. **264**(1): p. 110-6.

211. Tehei, M., et al., *Fast dynamics of halophilic malate dehydrogenase and BSA measured by neutron scattering under various solvent conditions influencing protein stability.* Proc Natl Acad Sci U S A, 2001. **98**(25): p. 14356-61.

212. Levandowsky, M., *Microbial life in extreme environments.* The Quarterly Review of Biology, 1979. **54**: p. 331-333.

213. D'Souza, S.E., W. Altekar, and S.F. D'Souza, *Adaptive response of Haloferax mediterranei to low concentrations of NaCl (< 20%) in the growth medium.* Arch Microbiol, 1997. **168**(1): p. 68-71.

214. Gupta, R., et al., *Microbial α-amylases: a biotechnological perspective* Process Biochemistry, 2003. **38**(11): p. 1599-1616.

215. van der Maarel, M.J., et al., *Properties and applications of starch-converting enzymes of the alpha-amylase family.* J Biotechnol, 2002. **94**(2): p. 137-55.

216. Stewart, E.J., F. Aslund, and J. Beckwith, *Disulfide bond formation in the Escherichia coli cytoplasm: an in vivo role reversal for the thioredoxins.* Embo J, 1998. **17**(19): p. 5543-50.

217. Ritz, D. and J. Beckwith, *Roles of thiol-redox pathways in bacteria.* Annu Rev Microbiol, 2001. **55**: p. 21-48.

218. Holmgren, A. and M. Bjornstedt, *Thioredoxin and thioredoxin reductase.* Methods Enzymol, 1995. **252**: p. 199-208.

219. Littlechild, J.A., J.E. Guy, and M.N. Isupov, *Hyperthermophilic dehydrogenase enzymes.* Biochem Soc Trans, 2004. **32**(Pt 2): p. 255-8.